(12) INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(19) World Intellectual Property Organization
International Bureau

(43) International Publication Date
21 May 2004 (21.05.2004)

PCT

(10) International Publication Number
WO 2004/042365 A2

(51) International Patent Classification[7]: G01N

(21) International Application Number:
PCT/US2003/036247

(22) International Filing Date:
3 November 2003 (03.11.2003)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:
60/423,113 1 November 2002 (01.11.2002) US

(71) Applicant *(for all designated States except US)*: **EVOLUTIONARY GENOMICS LLC** [US/US]; Bioscience Park Center, 12635 East Montview Boulevard, Suite 211, Aurora, CO 80010 (US).

(72) Inventor; and

(75) Inventor/Applicant *(for US only)*: **MESSIER, Walter** [US/US]; 2418 Bowen Street, Longmont, CO 80501 (US).

(74) Agents: **SWANSON, Barry, J.** et al.; Swanson & Bratschun, L.L.C., 1745 Shea Center Drive, Suite 330, Highlands Ranch, CO 80129 (US).

(81) Designated States *(national)*: AE, AG, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BW, BY, BZ, CA, CH, CN, CO, CR, CU, CZ, DE, DK, DM, DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW, MX, MZ, NI, NO, NZ, OM, PG, PH, PL, PT, RO, RU, SC, SD, SE, SG, SK, SL, SY, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VC, VN, YU, ZA, ZM, ZW.

(84) Designated States *(regional)*: ARIPO patent (BW, GH, GM, KE, LS, MW, MZ, SD, SL, SZ, TZ, UG, ZM, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, HU, IE, IT, LU, MC, NL, PT, RO, SE, SI, SK, TR), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, ML, MR, NE, SN, TD, TG).

Published:
— *without international search report and to be republished upon receipt of that report*

For two-letter codes and other abbreviations, refer to the "Guidance Notes on Codes and Abbreviations" appearing at the beginning of each regular issue of the PCT Gazette.

WO 2004/042365 A2

(54) Title: DEVELOPMENT OF THERAPEUTICS FOR THE TREATMENT OF ENDOTOXIN-MEDIATED DISEASES

(57) Abstract: The subject invention comprises a method for identifying an evolutionarily meaningful nucleotide change in a primate's *TLR4* polynucleotide. It further comprises methods for identifying agents that interact with the corresponding evolutionarily meaningful amino acid change so as to modulate the function of the TLR4 polypeptide, thereby attenuating activation of the NF-kB pathway. Such agents are useful in mitigating the LPS mediated response and in the treatment of sepsis, severe sepsis and septic shock.

DEVELOPMENT OF THERAPEUTICS FOR THE TREATMENT OF ENDOTOXIN-MEDIATED DISEASES

FIELD OF THE INVENTION

The present invention relates to methods to develop agents for treating endotoxin-mediated diseases of humans, such that the therapeutic agent interacts with the human TLR4 polypeptide extracellular domain in such a way as to attain a response resembling that of the Old World monkeys, specifically baboons and rhesus monkeys.

BACKGROUND OF THE INVENTION

Sepsis is a serious medical condition. According to R.A. Balk and L. C. Casey (April 2000 *Critical Care Clinics*):

- Sepsis results in 120,000 to 200,000 deaths annually in the United States
- Death due to this disease (4.2 deaths/100,000) has increased 82.6% from 1979 to 1997.
- It is the 12th leading cause of death overall and is the most common cause of shock encountered by internists in the U.S.
- Between 300,000 to 500,000 cases of sepsis are diagnosed per year
- Shock develops in about 40% of septic patients.
- Despite aggressive treatment, mortality ranges from 16% in patients with sepsis to 40-60% in patients with septic shock
- The annual health care cost from caring for patients with sepsis is $5-10 billion

To date, no truly effective therapy exists to counteract the effects of sepsis, although some techniques do show limited utility. An effective therapeutic approach would have tremendous social and commercial value. Described here is a method to develop to such a therapeutic.

Severe sepsis, also known as septic syndrome, refers to a chain of events leading from microbial infection to tissue injury and cardiovascular collapse. J.S. Stapczynski provides the following definitions:

- "Sepsis" is the systemic host response to infection, defined as SIRS (systemic inflammatory response syndrome) in combination with a documented infection
- "Severe sepsis" is defined as sepsis plus end-organ dysfunction or hypoperfusion
- "Septic shock" is defined as sepsis with hypotension, despite fluid resuscitation, and evidence of inadequate tissue perfusion

Significant complications from sepsis include central nervous system dysfunction, adult respiratory distress syndrome (ARDS), liver failure, acute renal failure (ARF), and disseminated intravascular coagulation (DIC). In different studies, the reported incidence rates of these complications in SIRS and sepsis is about 19% for CNS dysfunction, 2-8% for ARDS, 12% for liver failure, 9-23% for ARF, and 8-18% for DIC (Stapczynski, J.S. 2001 *eMedicine Journal* 2:7).

According to N.R. Chamberlain (2001 Bacterial sepsis with shock in *Infectious Diseases Lectures),* sepsis involves a very complex sequence of events and much remains incompletely understood about how a patient goes from SIRS to septic shock. Patients with septic shock have a biphasic immunological response. Initially they manifest an overwhelming inflammatory response to an infection. This is probably due to the pro-inflammatory cytokines tumor necrosis factor (TNF), IL-1, IL-12, Interferon gamma (IFNγ), and IL-6). The body then regulates this response by producing anti-inflammatory cytokines (IL-10), soluble inhibitors (TNF receptors, IL-1 receptor type II, and IL-1RA, an inactive form of IL-1). The patient manifests a period of immunodepression. Persistence of this hyporesponsiveness is associated with increased risk of nosocomial infection and death.

Approximately one half of septic shock cases are caused by Gram-negative bacteria (Balk, R.A., and Casey, L.C., April 2000 *Critical Care Clinics*). It has long been known that sepsis can be triggered by cell-wall components of Gram-negative bacteria, termed endotoxin (Takeda, K., and Akira, S. 2001 *Genes to Cells* 6:733-742).

Endotoxin is also associated with the development and progression of asthma, as well as other types of airway disease (Arbour, N.C. *et al.*, 2000 *Nature Genetics* 25:187-191). In asthma and airway disease, endotoxin is believed to influence

pathophysiological effects of air pollution (Arbour, N.C. *et al.*, 2000 *Nature Genetics* **25**:187-191). The incidence of asthma and airway diseases is increasing and, like sepsis, new treatments are needed. Effective therapeutics for these endotoxin-mediated diseases represent a serious unmet medical need.

Endotoxins are composed of a lipopolysaccharide (LPS) complex, containing Lipid A and polysaccharide. The TLR4 protein has been documented (Takeda, K., and Akira, S. 2001 *Genes to Cells* **6**:733-742) to recognize and bind LPS. This initiates a molecular cascade that triggers the innate immune system. Human TLR4 is known to be a homolog of the *Drosophila* protein Toll. Toll, like its human homolog, is necessary to initiate the innate immune response. Both Toll and TLR4 are known to signal through the NF-κB pathway (Medzhitov, R. *et al.*, 1997 *Nature* **388**:394-397).

One possible therapeutic avenue would involve inhibiting either the *TLR4* gene or, more likely, the TLR4 protein, (or perhaps administration of molecules that competitively inhibit TLR4). However, this is likely to have severe and undesirable side effects. Mice strains such as C3H/HeJ and C57BL/10ScCr are unresponsive to LPS, in contrast to wild type mice, as a result of genetic defects in *TLR4* (Rehli, M. *et al.*, 2000 *Journal of Biological Chemistry* **275**: 9773-9781). However, these strains are hypersensensitive to infection by Gram-negative bacteria (Beutler, B. 2002 *Current Opinion in Hematology* 9:2-10). Without a functional TLR4, and the innate immune response it triggers (which leads to an acquired immune response), these mice are unable to recognize these pathogenic invaders.

Most mammals are susceptible to septic shock. Humans, chimpanzees and bonobos are alike in extreme sensitivity to LPS, and to septic shock (Veloso, D. 1996 *Immunopharmacology* **33**(1-3): 374-376. However, it has been well established that baboons and rhesus monkeys are resistant to septic shock, even when confronted with very high levels of LPS (Redl, H. *et al.*, 1993 *Immunobiology* **187**:330-345; Veloso, D. 1996 *Immunopharmacology* **33**(1-3): 374-376). In fact, all the Old World monkeys may share this resistance to high levels of endotoxin-induced septic shock. Yet, in baboons and rhesus the innate immune response is known to be essentially the same as that of humans. Thus, baboons and rhesus have developed some mechanism for resistance to septic shock that does not interfere with innate immunity.

Because TLR4 protein is involved in septic shock, the inventors reasoned that differences in septic shock sensitivity between humans and baboons might be the

result of subtle differences in the TLR4 protein. Thus, information about the specific amino acid replacements that occurred during evolution could provide unparalleled insights into the mechanism by which baboons and rhesus monkeys resist LPS-induced septic shock while maintaining functional innate immunity.

Published *TLR4* sequences from human (GenBank AF177765, XM_057452, U88880, and U93091), bonobo (GenBank AF179220), and baboon (GenBank AF180964) were used to design primers for polymerase chain reaction (PCR) amplification of a set of *TLR4* homologs from various primates. The primate *TLR4* homologs that were amplified and sequenced included rhesus monkey, gorilla, chimpanzee, gibbon, squirrel monkey, and capuchin. In addition, *TLR4* was amplified and sequenced from human, bonobo, and baboon and the published sequences for these species were confirmed (Seq ID: 1 to 7). As noted in Table 1, in most cases only exons 2 and 3 were sequenced (these include the full coding region of the *TLR4* gene).

Table 1 *TLR4* Sequences

Seq. 1	Seq. 2	Seq. 3	Seq. 4	Seq. 5	Seq. 6	Seq. 7
Chimpanze e (Bonnie) *Exon 3*	**Chimpanze** e (Dustin) *Exons 2&3*	**Gorill** a *Exons 2&3*	**Gibbo** n *Exons 2&3*	**Rhesus monke** y *Exons 2&3*	**Capuchi** n *Exon 3*	**Squirre** l **monkey** *Exons 2&3*

These sequences were aligned and a series of molecular evolution analyses were then performed. Included in these analyses were Ka/Ks pairwise comparisons for each of these genes. Such pairwise comparisons calculate the differences between values of nonsynonymous nucleotide substitutions per nonsynonymous site (Ka) to synonymous substitutions per synonymous site (Ks). Ka values statistically significantly greater than the corresponding Ks values (Ka-Ks) strongly suggest the action of positive selection. Conversely, Ka values statistically significantly less than the corresponding Ks values (Ka-Ks) strongly suggest the action of negative selection,

i.e., evolutionary conservation. For convenience, these pairwise comparisons are most often displayed as ratios (Ka/Ks), such that Ka/Ks >1 signifies positive selection, while Ka/Ks <1 signifies conservation.

All of these coding sequence comparisons exhibited Ka/Ks ratios less than one, some with statistical significance. This is good evidence that these are generally well-conserved, which is a commonly observed pattern. However, even well-conserved proteins can have a limited number of amino acid changes in key domains that significantly affect the function of the protein. In such cases, Ka/Ks analysis of the entire coding sequence may indicate conservation, while Ka/Ks analysis of individual domain coding regions may indicate a positively selected domain within a conserved protein. Thus, polynucleotide sequences encoding individual domains of the TLR4 protein were also subjected to Ka/Ks analysis. Two key domains are an intracellular domain responsible for signaling and an extracellular domain responsible for LPS binding.

Intracellular signaling. Ka/Ks analysis of the polynucleotide coding sequence for the TIR domain, which is the intracellular domain of TLR4 protein responsible for signaling, and which initiates the NF-κB pathway, indicates that this domain is extremely well conserved. In fact, this analysis revealed some of the lowest Ka/Ks ratios ever documented. This indicates extreme evolutionary conservation, and strongly suggests two inferences. First, this domain is a crucial one, and generally cannot tolerate amino acid replacements. Second, the signaling pathway is likely to be unchanged in all these primates. That is, regardless of differences in LPS sensitivity, the cascade initiated by the TIR domain is likely to be biochemically similar in both humans and baboons. This result thus suggests that close attention must be paid to the extracellular domain of the TLR4 protein which governs LPS recognition.

Extracellular LPS binding. LPS is thought to bind to an extracellular domain. The extracellular binding domain of TLR4 includes a number of leucine-rich repeats (LRR). These are conserved between human, bonobo, and baboon, suggesting that the basic binding mechanism is unchanged between these species. In fact, the basic LRR structure is conserved even in the Toll homolog in *Drosophila*. However, Ka/Ks analysis performed on the LPS binding domain for each primate TLR4 protein indicated the baboon LPS binding domain may be positively selected relative to the human or bonobo LPS binding domain, although there was only one nonsynonymous

change, thus the result was not statistically significant. This is suggestive but inconclusive evidence that the difference in septic shock sensitivity between humans and baboons results from specific amino acid replacements in the LPS binding domain.

Ka/Ks analysis of the whole protein or critical domains did not provide conclusive information about the difference in sepsis susceptibility is humans *versus* baboons and rhesus monkeys, so we next looked at a specific amino acid of the *TLR4* gene. One human *TLR4* mutation (the "human null mutation") in the extracellular ligand binding domain has been reported (Arbour, N.C. *et al.*, 2000 *Nature Genetics* **25**:187-191) that results in complete lack of sensitivity to LPS. The *TLR4* gene from these individuals has been sequenced, and is available from GenBank (GenBank Acc. #1777766). Like baboons and rhesus, such individuals are resistant to septic shock (Arbour, N.C. *et al.*, 2000 *Nature Genetics* **25**:187-191). However, humans who are homozygous for this mutation have compromised immune systems and, like the C3H/HeJ and C57BL/10ScCr mice, LPS does not trigger innate immunity leaving them prone to serious Gram-negative bacterial infections.

The human null mutation is replacement of Asp299 by Gly299. Clearly, such a replacement results in substantial steric changes, leading to the loss of function observed in individuals with this mutation. Importantly, Asp299 is conserved in all mammalian TLR4s for which coding sequence data are available (except as noted below), even as phylogenetically distant from humans as mouse and rat. Such extensive conservation implies strong functional importance: this site does not generally tolerate amino acid substitutions. Importantly, however, we found that baboons and rhesus monkeys, and probably all Old World monkeys, do have an amino acid replacement at this site (Asp299 to Asn299).

	Amino acid 299	Septic Shock	Innate Immunity
Humans, most mammals, *Drosophila*	Asp	+	+
Human null mutation	Gly	-	-
Old World monkeys (Rhesus, baboons, etc.)	Asn	-	+

It is clear that inhibiting TLR4 function completely is not a viable therapeutic approach because it results in too great an impairment of the immune response.

Similarly, modeling a therapeutic based on the human TLR4 Gly299 mutation also would result in susceptibility to Gram-negative bacterial infections. In contrast, baboons both resist septic shock and are fully capable of recognizing and addressing Gram-negative bacterial infections. This leads to a novel approach for developing therapeutic agents for treatment of endotoxin-mediated diseases of humans.

SUMMARY OF THE INVENTION

The subject invention comprises a method of identifying an evolutionarily meaningful nucleotide change in a first primate's polynucleotide wherein the first primate's polynucleotide may be associated with a physiological condition that is present or enhanced in the first primate relative to a second primate, comprising the steps of: (a) comparing polynucleotide sequences of the first primate with corresponding polynucleotide sequences of the second primate to identify a polynucleotide that has been overall negatively selected in the first and second primates; and (b) identifying in the first primate's overall negatively selected polynucleotide, an evolutionarily meaningful nucleotide change, whereby the nucleotide change in the first primate's negatively selected polynucleotide may be associated with the physiological condition in the first primate.

The evolutionarily meaningful nucleotide change is a nonsynonymous nucleotide change in an otherwise conserved polynucleotide that is or is believed to be associated with a physiological condition. The analysis of the polynucleotides to determine whether they are negatively selected or conserved can be carried out by any method known in the art, but preferably is accomplished by a KA/KS-type analysis as described herein.

A nucleotide change in a primate's negatively selected polynucleotide can be correlated with a physiological condition in the primate by analyzing the functional effect of the presence or absence of the identified nonsynonymous nucleotide change in a model *in vivo*, *ex vivo* or *in vitro* system using methods known in the art.

In one embodiment, the first primate is an Old World monkey, the second primate is Homo sapiens, and the negatively selected polynucleotide is TLR4 polynucleotide. The evolutionarily meaningful nucleotide change in the TLR4 polynucleotide results in an Asp299 in the human and an Asn299 in the Old World monkey, e.g., baboon and rhesus monkey.

The present invention also relates to methods to develop agents for treating endotoxin-mediated diseases of humans, such that the therapeutic agent interacts with the human TLR4 polypeptide extracellular domain in such a way as to attain a response resembling that of the Old World monkeys, specifically baboons and rhesus monkeys.

It has been suggested (Arbour, N.C. *et al.*, 2000 *Nature Genetics* **25**:187-191) that the region of the TLR4 receptor around residue 299 has an α-helical structure. The substitution of the glycine residue (as found in the human null mutant) for the aspartic acid residue likely disrupts the 3-D structure of this helix in a catastrophic manner. However, the asparagine residue found at position 299 in the baboon and rhesus sequences is a biochemically conservative replacement compatible with the helical structure. This evolutionarily tolerated, structurally-conservative replacement allows baboons and rhesus monkeys (and most likely, all the Old World monkeys) to modulate the interaction with LPS, such that Gram-negative bacteria still trigger the innate immune response in such a way that the known resistance of both baboon and rhesus to extremely high levels of LPS is achieved. An alignment of TLR4 protein sequences for the region of the protein that flanks this residue (from a number of mammalian species) is shown in Figure 1.

It is believed that the Asp to Asn amino acid replacement at position 299 confers resistance to septic shock in baboons. Therefore, transgenic mice whose *TLR4* encodes the Asn replacement, when compared to controls (transfected with 'normal' mouse *TLR4*), should show that the experimental transgenics exhibit an increased resistance to septic shock, i.e., they will tolerate much higher levels of LPS and/or live bacteria.

The insight of the invention described herein is that the Asp299 residue is critical to initiation of the LPS-triggered cascade that leads to endotoxin-mediated diseases, such as septic shock and asthma and other airway disorders. In one preferred embodiment, a therapeutic agent is developed that, when administered, causes human TLR4 to react to exposure to endotoxin in the way baboon or rhesus monkey TLR4 does. Accordingly, in one aspect, a method is provided whereby a peptide therapeutic agent could be isolated. Such a peptide would reduce access by LPS to the key amino acid of TLR4 determining septic shock, Asp299. If delivered during an episode of acute septic shock, the peptide should "derail" the cascade that is initiated when LPS and/or live bacteria encounters the human TLR4 protein. Such a

peptide agent can be easily tested in rodent models. Successful demonstration of protection of such models from septic shock would pave the way to similar human trials of such peptide agent. Because such a peptide would only be administered during acute episodes of septic shock, possible problems stemming from repeated administration and subsequent sensitization would be minimized. Those skilled in the art can easily determine the optimal length and amino acid composition of such therapeutic peptide, which can be further refined by testing in rodent models, using methods known in the art.

In another embodiment, a therapeutic peptide could be designed that had the same sequence as the region surrounding Asp299. Such a therapeutic could be useful as a decoy to bind to LPS and reduce the amount of LPS available to bind to the TLR4 protein.

In another embodiment, an antibody or portion of an antibody could be isolated or designed that could attenuate access by endotoxin to Asp299 such that the endotoxin-mediated cascade is reduced. In a preferred embodiment, the antibody or fragment thereof is directed to an epitope that includes the Asp299 residue; the epitope preferably is an amino acid segment of 10 or less residues containing the Asp299 residue.

In another embodiment, a small molecule is identified that will reduce access to the critical Asp299 residue, or modulate the interaction of endotoxin with the Asp299 residue such that the cascade leading to endotoxin-mediated disease is modulated.

Another embodiment is to use the method disclosed herein to develop a therapeutic agent to treat human asthma. Small molecules could be designed or identified by screening libraries of small molecules that interact with Asp299 of the TLR4 polypeptide or the region containing the Asp299 polypeptide. Such therapeutic agents could be used to ameliorate the severity of asthmatic episodes. It is also likely that some therapeutic agents identified using the methods of this invention could be administered on a regular basis to reduce the effects of chronic airway diseases.

Using the teachings provided herein, persons skilled in the art will recognize that therapeutics can be developed for other diseases involving LPS – TLR4 protein interactions.

All references cited herein are each incorporated herein by reference in their entirety.

BRIEF DESCRIPTION OF THE DRAWINGS

Figure 1 is an alignment of TLR4 protein sequences for the region of the protein that flanks the Asp299 residue from a number of mammalian species. Amino acid residues are shown in the single letter IUPAC code. Residues that are identical in all species examined are shown in bold. Dashes have been introduced (where insertions or deletions have occurred) to maximize the alignment. The critical residue (human Asp299, baboon Asn299) is shown in lower case. Note that this Asp residue is conserved in all mammal species examined, with the exception of the biochemically-conservative Asn replacement in the Old World monkeys baboon and rhesus (and, importantly, the non-functional human null mutant).

Figure 2 is the nucleotide sequence for a first chimpanzee's ("Bonnie") *TLR4* exon 3.

Figure 3 is the nucleotide sequence for a second chimpanzee's ("Dustin") *TLR4* exons 2 and 3.

Figure 4 is the nucleotide sequence for gorilla *TLR4* exons 2 and 3.

Figure 5 is the nucleotide sequence for gibbon *TLR4* exons 2 and 3.

Figure 6 is the nucleotide sequence for rhesus monkey *TLR4* exons 2 and 3.

Figure 7 is the nucleotide sequence for capuchin *TLR4* exon 3.

Figure 8 is the nucleotide sequence for squirrel monkey *TLR4* exon 3.

DETAILED DESCRIPTION OF THE INVENTION

The subject invention relates to a method of identifying a nucleotide change in a TLR4 polynucleotide sequence of an Old World monkey wherein such change may be associated with reduced sensitivity to Gram-negative bacterial infection. This method involves the comparison of the TLR4 polynucleotide sequence from the Old World monkey with corresponding TLR4 polynucleotide sequence of a human to identify a polynucleotide change in said Old World monkey's TLR4 sequence that is evolutionarily meaningful. The evolutionarily meaningful change may then be associated with reduced sensitivity to Gram-negative bacterial infection. In particular, the evolutionarily meaningful change is from Asp299 in the human to Asn299 in the rhesus monkey or baboon.

The subject invention also includes a method of identifying a therapeutic agent that reduces sensitivity to Gram-negative bacterial infection. This method comprises

(a) contacting candidate agents with human TLR4 polypeptide; and (b) identifying a therapeutic agent that interacts with the TLR4 polypeptide to substantially reduce sensitivity to Gram-negative bacterial infection. The agent interaction with TLR4 polypeptide preferably occurs at Asp299.

The therapeutic agent identified according to the subject invention can be used in the treatment of sepsis, severe sepsis, septic shock, asthma or other respiratory ailments, in humans or non-human primates.

The subject invention also concerns a number of novel *TLR4* polynucleotide sequences as provided in Figures 2-9, and their deduced TLR4 polypeptide sequences. These sequences are chimpanzee, gorilla, gibbon, rhesus monkey, capuchin, squirrel monkey and baboon. They are useful in the identification of evolutionarily meaningful nucleotide changes in other primate *TLR4* polynucleotides. Their polynucleotide or polypeptide sequences may also be useful in the design of candidate therapeutic agents according to the subject invention.

The practice of the present invention employs, unless otherwise indicated, conventional techniques of molecular biology, genetics, and assay development, which are within the skill of the art. Such techniques are explained fully in the literature, such as : "Molecular Cloning: A Laboratory Manual", second edition (Sambrook et al., 1989); "Oligonucleotide Synthesis" (M.J. Gait, ed., 1984); "Current Protocols in Molecular Biology" (F.M. Ausubel et al., eds., 1987); "PCR: The Polymerase Chain Reaction", (Mullis et al., eds., 1994)

Definitions

As used herein, a "polynucleotide" refers to a polymeric form of nucleotides of any length, either ribonucleotides or deoxyribonucleotides, or analogs thereof. This term refers to the primary structure of the molecule, and thus includes double- and single-stranded DNA, as well as double- and single-stranded RNA. It also includes modified polynucleotides such as methylated and/or capped polynucleotides. The terms "polynucleotide" and "nucleotide sequence" are used interchangeably.

As used herein, a "gene" refers to a polynucleotide or portion of a polynucleotide comprising a sequence that encodes a protein. It is well understood in the art that a gene also comprises non-coding sequences, such as 5' and 3' flanking sequences (such as promoters, enhancers, repressors, and other regulatory sequences) as well as introns.

The terms "polypeptide," "peptide," and "protein" are used interchangeably herein to refer to polymers of amino acids of any length. These terms also include proteins that are post-translationally modified through reactions that include glycosylation, acetylation and phosphorylation.

The term "K_A/K_S-type methods" means methods that evaluate differences, frequently (but not always) shown as a ratio, between the number of nonsynonymous substitutions and synonymous substitutions in homologous genes (including the more rigorous methods that determine non-synonymous and synonymous sites). These methods are designated using several systems of nomenclature, including but not limited to K_A/K_S, d_N/d_S, D_N/D_S.

The term "evolutionarily meaningful change" refers to one or more nonsynonymous nucleotide change(s) or corresponding amino acid change(s) between two species that occurs in an otherwise conserved polynucleotide or polypeptide, that may be attributed to a positive selective pressure, and which is or is believed to be associated with a physiological condition. A conserved polynucleotide can be identified by methods known in the art including a K_A/K_S-type analytical method. Typically, a K_A/K_S ratio less than about 1.0, more preferably less than about 0.75, and most preferably less than about 0.5 indicates the action of negative selection. The presence of a nonsynonymous nucleotide change in such a conserved polynucleotide (i.e., containing no other nucleotide changes or only synonymous nucleotide changes) is considered to be an evolutionarily meaningful change. The phrase "associated with a physiological condition" means that the nonsynonymous nucleotide change has been observed in individuals to result in the physiological condition at issue, has been shown to be involved in a molecular mechanism related to the physiological condition, and/or occurs in a location of the gene that is relevant to a protein function that is essential to the occurrence of the physiological condition. For example, as discussed herein, a nonsynonymous nucleotide change in the baboon or rhesus monkey gene is believed to be associated with the physiological condition of enhanced resistance to the endotoxin-mediated response.

The term "positive evolutionarily meaningful change" means an evolutionarily meaningful change in a particular species that results in an adaptive change that is positive as compared to other related species. An example of a positive evolutionarily meaningful change is a change that has resulted in reduced sensitivity to the LPS mediated response.

The term "resistant" means that an organism exhibits an ability to avoid, or diminish the extent of, a disease condition and/or development of the disease, preferably when compared to non-resistant organisms.

The term "susceptibility" means that an organism fails to avoid, or diminish the extent of, a disease condition and/or development of the disease condition, preferably when compared to an organism that is known to be resistant.

It is understood that resistance and susceptibility vary from individual to individual, and that, for purposes of this invention, these terms also apply to a group of individuals within a species, and comparisons of resistance and susceptibility generally refer to overall, average differences between species, although intra-specific comparisons may be used.

The term "nucleotide change" refers to nucleotide substitution, deletion, and/or insertion, as is well understood in the art.

The term "agent", as used herein, means a biological or chemical compound such as a simple or complex organic or inorganic molecule, a peptide, a protein or an oligonucleotide that modulates the function of a polypeptide. A vast array of compounds can be synthesized, for example oligomers, such as oligopeptides and oligonucleotides, and synthetic organic and inorganic compounds based on various core structures, and these are also included in the term "agent". In addition, various natural sources can provide compounds for screening, such as plant or animal extracts, and the like. The term "agent" can include or exclude antibodies or fragments thereof. Compounds can be tested singly or in combination with one another.

The term "therapeutic agent" as used herein means an agent as described above used to treat a disease or condition.

The term "to modulate function" of a polypeptide means that the function of the polypeptide is altered in the presence of an agent compared to the absence of the agent. Modulation may occur on any level that affects function. Modulation of a polypeptide function may be direct or indirect, and measured directly or indirectly.

A "function of a polypeptide" includes, but is not limited to, conformation, folding (or other physical characteristics), binding to other moieties (such as ligands), activity (or other functional characteristics), and/or other aspects of protein structure or functions. For example, an agent that acts on a polypeptide and affects its conformation, folding (or other physical characteristics), binding to other moieties

(such as ligands), activity (or other functional characteristics), and/or other aspects of protein structure or functions is considered to have modulated polypeptide function. The ways that an effective agent can act to modulate the function of a polypeptide include, but are not limited to 1) changing the conformation, folding or other physical characteristics; 2) changing the binding strength to its natural ligand or changing the specificity of binding to ligands; and 3) altering the activity of the polypeptide.

The term "to modulate the endotoxin or LPS mediated response" means that the function of the TLR 4 polypeptide is altered in the presence of an agent compared to the absence of the agent. The modulation reduces the clinical symptoms of sepsis, severe sepsis, or septic shock, including central nervous system dysfunction, adult respiratory distress syndrome, liver failure, acute renal failure, disseminated intravascular coagulation, and the like. Preferably, these symptoms are reduced in increasing preference by at least about 10%, 20%, 30%, 40%, 50%, 60%, 70%, 80%, 90% and 100%, as measured by standard clinical indicators or assays for those symptoms as is known in the art. Modulation can be detected in other ways as well, including reduced affinity or altered kinetics of LPS binding to the TLR4 extracellular domain, or reduced signaling to a downstream effector of TLR4 in the LPS mediated response. Preferably, the agent modulation results in the attenuation of the LPS mediated response to the degree that the human TLR4 interacts with the LPS in the same manner as the Old World monkey TLR4.

Further, "modulation of endotoxin or LPS mediated response" means significant reduction or attenuation of the response whereby the clinical symptoms of sepsis, severe sepsis or septic shock are reduced as indicated above. It does not refer to abrogation or elimination of the response; i.e., innate immunity remains intact. The therapeutic agent of the subject invention is one which interacts directly or indirectly with the Asp299 residue such that the activation of the NF-kB pathway and the clinical symptoms of sepsis, severe sepsis and septic shock are attenuated.

The term "target site" means a location in a polypeptide which can be a single amino acid and/or is a part of, a structural and/or functional motif, e.g., a binding site, a dimerization domain, or a catalytic active site. Target sites may be useful for direct or indirect interaction with an agent, such as a therapeutic agent.

The term "positively selected" means an evolutionarily significant change in a particular organism, species, subspecies, variety, cultivar or strain that results in an adaptive change that is positive as compared to other related organisms. An example

of a positive evolutionarily significant change is a change that has resulted in enhanced yield in crop plants. As stated above, positive selection is indicated by a K_A/K_S ratio greater than 1.0. With increasing preference, the K_A/K_S value is greater than 1.25, 1.5 and 2.0.

"TLR4" as used herein refers to the polynucleotide encoding the Toll like receptor 4 polypeptide.

"TLR4" as used herein refers to the polypeptide encoded by *TLR4* polynucleotide.

"LPS" as used herein refers to lipopolysaccharide and is used interchangeably with the word "endotoxin".

General Methods of the Invention

The general method of the invention is as follows. Briefly, the *TLR4* polynucleotide sequences are obtained from a human source and a number of non-human primate sources. They are compared to one another to determine whether the TLR4 polynucleotide is conserved or negatively selected. Then, having determined that the polynucleotide is conserved, the *TLR4* polynucleotide sequences from the human and other primates are analyzed to identify any nonsynonymous or evolutionarily meaningful nucleic acid differences. The *TLR4* sequences from each species are then characterized in terms of whether they do or do not correlate with decreased sensitivity to Gram-negative bacterial infection for that species, thereby indicating those evolutionarily meaningful changes that could be or are associated with the decreased sensitivity to LPS. This method resulted in the identification of Asn299 as the critical amino acid found in baboons and rhesus monkeys that confers attenuated sensitivity to LPS in those species as compared to humans which have Asp299.

U.S. Serial No. 10/100,422, filed March 18, 2002, and incorporated herein in its entirety by reference, describes in detail a number of methods useful in sequencing homologous polynucleotide sequences from primates, and methods for identification of evolutionarily meaningful changes in polynucleotides that can be correlated with a particular physiological condition in humans or in non-human primates.

Determination of evolutionarily meaningful changes first requires a determination as to whether the polynucleotide at issue is negatively selected. Any of several different molecular evolution analyses or K_A/K_S-type methods can be

employed to determine whether the human gene sequences and the non-human primate polynucleotide are conserved or negatively selected. Kreitman and Akashi (1995) *Annu. Rev. Ecol. Syst.* 26:403-422; Li, *Molecular Evolution*, Sinauer Associates, Sunderland, MA, 1997. For example, negative selection on proteins (*i.e.*, molecular-level conservation) can be detected in protein-coding genes by pairwise comparisons of the ratios of nonsynonymous nucleotide substitutions per nonsynonymous site (K_A) to synonymous substitutions per synonymous site (K_S) (Li *et al.*, 1985; Li, 1993). Any comparison of K_A and K_S may be used, although it is particularly convenient and most effective to compare these two variables as a ratio. Negatively selected polynucleotides are identified by having a K_A/K_S ratio of less than 1.0, preferably less than 0.75 and more preferably less than 0.5. Preferably, the K_A/K_S analysis by Li *et al.* is used to carry out the present invention, although other analysis programs that can detect negatively selected genes between species can also be used. Li *et al.* (1985) *Mol. Biol. Evol.* 2:150-174; Li (1993); see also *J. Mol. Evol.* 36:96-99; Messier and Stewart (1997) *Nature* 385:151-154; Nei (1987) *Molecular Evolutionary Genetics* (New York, Columbia University Press). The K_A/K_S method, which comprises a comparison of the rate of non-synonymous substitutions per non-synonymous site with the rate of synonymous substitutions per synonymous site between homologous protein-coding region of genes in terms of a ratio, can be used to identify conserved polynucleotides. A synonymous ("silent") substitution is one that, owing to the degeneracy of the genetic code, makes no change to the amino acid sequence encoded; a nonsynonymous substitution results in an amino acid replacement. The extent of each type of change can be estimated as K_A and K_S, respectively, the numbers of synonymous substitutions per synonymous site and non-synonymous substitutions per non-synonymous site. Calculations of K_A/K_S may be performed manually or by using software. An example of a suitable program is MEGA (Molecular Genetics Institute, Pennsylvania State University).

For the purpose of estimating K_A and K_S, either complete or partial human and primate protein-coding sequences are used to calculate total numbers of synonymous and non-synonymous substitutions, as well as non-synonymous and synonymous sites. The length of the polynucleotide sequence analyzed can be any appropriate length, but is preferably at least 60 nucleotides in length. Preferably, the entire coding sequence is compared, in order to determine overall conservation. Publicly available

computer programs, such as Li93 (Li (1993) *J. Mol. Evol.* 36:96-99) or INA, can be used to calculate the K_A and K_S values for all pairwise comparisons.

As indicated above, conservation is indicated by the K_A/K_S ratio being less than about 1.0, more preferably less than about 0.75, and most preferably less than about 0.5. Preferably, statistical analysis is performed on all decreased K_A/K_S ratios, including, but not limited to, standard methods such as Student's *t*-test and likelihood ratio tests described by Yang (1998) *Mol. Biol Evol.* 37:441-456.

All methods for calculating K_A/K_S ratios are based on a pairwise comparison of the number of nonsynonymous substitutions per nonsynonymous site to the number of synonymous substitutions per synonymous site for the protein-coding regions of homologous genes from related species. Each method implements different corrections for estimating "multiple hits" (*i.e.,* more than one nucleotide substitution at the same site). Each method also uses different models for how DNA sequences change over evolutionary time. Thus, preferably, a combination of results from different algorithms is used to increase the level of sensitivity for detection of negatively-selected genes and confidence in the result.

As discussed above, the foregoing methods resulted in the identification of an evolutionarily meaningful nucleotide change in the conserved *TLR4* polynucleotide. The Asn299 of baboons and rhesus monkeys was found to attenuate LPS sensitivity in those species relative to humans which have Asp299. This information can be useful in the development of agents that interact with human TLR4 Asp299 in such a manner so as to attenuate activation of the NF-kB pathway by LPS, therefore aiding in the treatment of sepsis, severe sepsis and septic shock.

An agent is designed or identified that will interact with human TLR4 protein in such a way that the agent modulates the human endotoxin mediated response. Preferably, the agent causes the human TLR4 to interact with LPS in a manner that is similar to that of baboon or rhesus TLR4 interaction with LPS. Such agents can be peptide, protein, organic molecules, or aptamers, or whatever agent can have the specific effect.

Generally, techniques of combinatorial chemistry can be used to generate numerous permutations of agent candidates to be screened for effectiveness in reducing access of LPS to Asp299 of TLR4. Those of skill in the art can devise and/or obtain suitable agents for testing. In general, screening can be performed by adding an agent to a sample of isolated TLR4 or appropriate cells expressing TLR4

and monitoring the effect, that is, modulation of the cascade known to lead to endotoxin-mediated disease. The experiments preferably include a control sample which does not receive the candidate agent. Differences between treated and untreated cells indicate effects attributable to the candidate agent. Optimally, a greater effect is seen in the presence of the candidate agent than in the absence of the candidate agent. Hoshino, K., Takeuchi, O., Kawai, T., et al. (1999). *J. Immunol.* 162, 3749-3752. describe a typical assay for measuring the NF-kB pathway that is known to lead to endotoxin-mediated disease.

The screening methods for agents that interact with TLR4 polypeptide can be carried out *in vitro*, *ex vivo* or *in vivo* using TLR4 protein or polypeptide or extracellular fragment thereof, or NF-kB pathway models known in the art.

In an example for an assay for an agent that binds to TLR4 polypeptide, an affinity column is prepared with purified human TLR4 or a synthetically prepared peptide of a small region of TLR4 containing Asp299. The affinity column is then used to screen a library of compounds (libraries of compounds include, but are not limited to, peptides, aptamers, small molecules, etc.) which have been appropriately labeled. Suitable labels include, but are not limited to fluorochromes, radioisotopes, enzymes and chemiluminescent compounds. The unbound and bound compounds can be separated by washes using various conditions known to those skilled in the art. In addition to affinity columns, there are other techniques, such as measuring the fluorescence anisotropy of a protein which will change upon binding another molecule. For example, a BIAcore assay using a sensor chip (supplied by Pharmacia Biosensor, Stitt et al. (1995) Cell 80:661-670.) that is coupled to TLR4 or a peptide of TLR4 containing Asp299 may be performed to determine the binding activity of different agents.

It is also understood that the in vitro screening methods of this invention include structural, or rational, drug design, in which the amino acid sequence surrounding Asp299, three-dimensional atomic structure or other property (or properties) of the amino acid sequence surrounding Asp299 provides a basis for designing an agent which is expected to bind to TLR4 in such a way as to reduce access by endotoxin to Asp299.

The screening methods described above represent primary screens, designed to detect any agent that may bind to TLR4 and/or exhibit activity that modulates the function of TLR4. The skilled artisan will recognize that secondary tests will likely

be necessary in order to evaluate an agent further. For example, a secondary screen may comprise testing the agent(s) in a mouse model or other animal models for effect in reducing the endotoxin-mediated cascade leading to disease.

The invention also includes agents identified by the screening methods described herein.

Peptide agents

A peptide agent can be isolated by screening a library of randomly synthesized peptides for peptides that bind to residue 299, using methods known in the art, for example, as described in Dower, William J; Cwirla, Steven E; Barrett, Ronald W, "Peptide library and screening systems." *Biotechnol Advances* 1996 14(4):490. Peptides found to bind to TLR4 in such a way that the endotoxin does not interact directly with Asp 299 are selected. In one embodiment, the peptide library contains randomly synthesized peptides of at least 20 amino acids in length, preferably less than 50 amino acids in length. In another embodiment, the peptide library contains randomly synthesized peptides of between 15 and 20 amino acids in length. In a third embodiment, the peptide library contains peptides between 10 and 14 amino acids in length. In a fourth embodiment, the peptide library contains peptides between 1 and 9 amino acids in length. Peptides found to bind TLR4 at or near residue 299 are then subjected to secondary screens in vitro for effect on the binding of LPS to TLR4 and prevention or attenuation of TLR4 signaling.

Aptamers

Similar to that described above, the subject invention includes methods to identify an aptamer agent by screening a library of randomly synthesized single-stranded nucleotides, using methods known in the art, for example, as described in Bell, C.; Lynam, E.; Landfair, D.J.; Janjic, N.; and Wiles, M.E., "Oligonucleotides NX1838 inhibits VEGF165-mediated cellular responses in vitro", *In Vitro Cell Dev. Biol. Anim.* 1999 Oct:35(9):533-42.

Small molecule agents

One method for developing such a small molecule agent would be to use 3-dimensional modeling of the secondary and tertiary structure of the region of TLR4 that surrounds the critical Asp299 residue. Potential small molecules can be 'docked'

in silico, in order to identify a close fit that will inhibit access to the aspartic acid residue or have the desired structural affect. Likely candidates can then be tested *in vivo* in rodent models. Small molecule therapeutics could also be identified from libraries of small molecules through the use of assays that screen for compounds that bind to the region of TLR4 containing Asp299. In a similar fashion, it may be possible to design or screen for a small molecule that does not completely block access to the Asp299 residue, but rather modulates the kinetics of LPS binding in this region in such a way that it more closely resembles the kinetics of LPS binding to baboon TLR4 protein.

Therapeutic compositions that comprise agents

As described herein, agents can be screened for their capacity to modulate the LPS mediated NF-kB pathway.

Various delivery systems are known in the art that can be used to administer agents identified according to the subject methods. Such delivery systems include aqueous solutions, encapsulation in liposomes, microparticles or microcapsules or conjugation to a moiety that facilitates intracellular admission.

Therapeutic compositions comprising agents may be administered parenterally by injection, although other effective administration forms, such as intra-articular injection, inhalant mists, orally-active formulations, transdermal iontophoresis or suppositories are also envisioned. The carrier may contain other pharmacologically-acceptable excipients for modifying or maintaining the pH, osmolarity, viscosity, clarify, color, sterility, stability, rate of dissolution, or odor of the formulation. The carrier may also contain other pharmacologically-acceptable excipients for modifying or maintaining the stability, rate of dissolution, release or absorption of the agent. Such excipients are those substances usually and customarily employed to formulate dosages for parenteral administration in either unit dose or multi-dose form.

Once the therapeutic composition has been formulated, it may be stored in sterile vials as a solution, suspension, gel, emulsion, solid, or dehydrated or lyophilized powder. Such formulations may be stored either in a ready to use form or requiring reconstitution immediately prior to administration. The manner of administering formulations containing agents for systemic delivery may be via subcutaneous, intramuscular, intravenous, intranasal or vaginal or rectal suppository.

The amount of agent which will be effective in the treatment of a particular disorder or condition will depend on the nature of the disorder or condition, which can be determined by standard clinical techniques. In addition, *in vitro* or *in vivo* assays may optionally be employed to help identify optimal dosage ranges. The precise dose to be employed in the formulation will also depend on the route of administration, and the seriousness or advancement of the disease or condition, and should be decided according to the practitioner and each patient's circumstances. Effective doses may be extrapolated from dose-response curves derived from *in vitro* or animal model test systems. For example, an effective amount of an agent identified according to the subject methods is readily determined by administering graded doses of the agent and observing the desired effect.

The following examples are provided to further assist those of ordinary skill in the art. Such examples are intended to be illustrative and therefore should not be regarded as limiting the invention. A number of exemplary modifications and variations are described in this application and others will become apparent to those of skill in this art. Such variations are considered to fall within the scope of the invention as described and claimed herein.

Example 1. PCR amplification and DNA sequencing of primate *TLR4* sequences.

Published *TLR4* sequences from human (GenBank AF177765, XM_057452, U88880, and U93091), bonobo (GenBank AF179220), and baboon (GenBank AF180964) were used to design primers (by methods well-known to those skilled in the art) for polymerase chain reaction (PCR) amplification of a set of *TLR4* homologs from various primates. The primate *TLR4* homologs that were PCR amplified and DNA sequenced (by methods well-known to those skilled in the art) included rhesus monkey, gorilla, chimpanzee, gibbon, squirrel monkey, and capuchin. In addition, *TLR4* was amplified and sequenced from human, bonobo, and baboon and the published sequences for these species were confirmed (Seq ID: 1 to 7). Because exons 2 and 3 contain the full coding region of the *TLR4* gene, in most cases only exons 2 and 3 were sequenced. These sequences were aligned by methods well-known to those skilled in the art.

Example 2. Ka/Ks analysis.

Ka/Ks pairwise comparisons were completed for each of these genes. Such pairwise comparisons calculate the differences between values of nonsynonymous nucleotide substitutions per nonsynonymous site (Ka) to synonymous substitutions per synonymous site (Ks). Ka values statistically significantly greater than the corresponding Ks values (Ka-Ks) strongly suggest the action of positive selection. Conversely, Ka values statistically significantly less than the corresponding Ks values (Ka-Ks) strongly suggest the action of negative selection, i.e., evolutionary conservation. For convenience, these pairwise comparisons are most often displayed as ratios (Ka/Ks), such that Ka/Ks >1 signifies positive selection, while Ka/Ks <1 signifies conservation.

All of these whole protein comparisons exhibited Ka/Ks ratios less than one, some with statistical significance. This is good evidence that these are generally well-conserved proteins, which is a commonly observed pattern. However, even well-conserved proteins can have amino acid changes in key domains that significantly affect the function of the protein. In such cases, Ka/Ks analysis of the entire coding sequence may indicate conservation, while Ka/Ks analysis of individual domain coding regions may indicate a positively selected domain within a conserved protein. Thus, polynucleotides encoding individual domains of the TLR4 protein were also subjected to analysis. Two key domains are an intracellular domain responsible for signaling and an extracellular domain responsible for LPS binding. Ka/Ks analysis was performed for the TIR domain, which is the intracellular domain of TLR4 protein responsible for signaling, and which initiates the NF-κB pathway. This analysis indicated that this domain is extremely well conserved. In fact, this analysis revealed some of the lowest Ka/Ks ratios ever documented. Ka/Ks analysis was then performed for the extracellular signaling domain of TLR. Here the result was inconclusive, in that although evidence was seen for possible positive selection on the extracellular LPS-binding domain of baboon TLR4 (relative to human TLR4), no statistical support exists for this. As a result, further analysis was performed (see Example 3).

Example 3. Further molecular evolutionary analysis.

Further analysis included a search for individual amino acid replacements that are either shared by, or are unique to, the human and baboon *TLR4* sequences. One

human *TLR4* mutation in the extracellular ligand binding domain has been reported (Arbour, N.C. *et al.*, 2000 *Nature Genetics* **25**:187-191) that results in complete lack of sensitivity to LPS. Like baboons and rhesus, such individuals are resistant to septic shock. However, humans who are homozygous for this mutation have compromised immune systems and LPS does not trigger innate immunity leaving them prone to serious Gram-negative bacterial infections. The human null mutation is replacement of Asp299 by Gly299. Importantly, Asp299 is conserved in all mammalian *TLR4*s for which sequence is available, even as phylogenetically distant as mouse and rat, with the exception of baboon and rhesus, which have a biochemically conservative replacement amino acid replacement at this site (Asp299 to Asn299). The substitution of the glycine residue (as found in the human null mutant) for the aspartic acid residue likely disrupts the 3-D structure of this helix in a catastrophic manner. However, because the asparagine residue found at position 299 in the baboon and rhesus sequences is a biochemically conservative replacement, it is likely compatible with the helical structure. This evolutionarily tolerated, structurally-conservative replacement thus probably allows baboons and rhesus monkeys (and most perhaps, all the Old World monkeys) to modulate the interaction with LPS, such that Gram-negative bacteria still trigger the innate immune response in such a way that the known resistance of both baboon and rhesus to extremely high levels of LPS is achieved.

Example 4. Design of a peptide therapeutic agent

A library of random peptides 20 amino acids long is synthesized and screened for peptide agents that bind to TLR4. A secondary screen then assesses peptides found in this primary screen for ability to reduce access by LPS to Asp299 and measurably reducing the LPS-mediated cascade leading to septic shock. The optimal length and amino acid composition of the therapeutic peptide agent can be refined by testing in rodent models, as would be known to one skilled in the art.

Example 5. Design and synthesis of a decoy peptide agent.

A peptide is designed that has the same sequence as the region of human TLR4 from amino acid 289 through 309. This peptide is synthesized synthetically and formulated for delivery as a therapeutic. Such a peptide therapeutic would be useful

as a decoy to bind to LPS and reduce the amount of LPS available to bind to the TLR4 protein.

Example 6. Design and synthesis of a small molecule therapeutic agent.

The secondary and tertiary structure of the region of TLR4 protein from amino acid 289 through 309 is modeled. Small molecule core structures are screened *in silico* for the ability to 'dock' in this region in order to identify a close fit that will inhibit access by endotoxin to the Asp299 residue either by directly or indirectly reducing access to Asp299. Likely candidates can then be tested *in vivo* in rodent models.

Example 7. Screening for a small molecule therapeutic.

Libraries of small molecules are purchased from one of several vendors. An affinity column is prepared with purified human TLR4 or a synthetically prepared peptide of a small region of TLR4 containing Asp299. The affinity column is then used to screen a library of compounds which have been appropriately labeled. Suitable labels include, but are not limited to fluorochromes, radioisotopes, enzymes and chemiluminescent compounds. The unbound and bound compounds can be separated by washes using various conditions known to those skilled in the art. In addition to affinity columns, there are other techniques, such as measuring the fluorescence anisotropy of a protein which will change upon binding another molecule. For example, a BIAcore assay using a sensor chip (supplied by Pharmacia Biosensor, Stitt et al. (1995) Cell 80:661-670.) that is coupled to TLR4 or a peptide of TLR4 containing Asp299 may be performed to determine the binding activity of different agents.

A secondary screen is then employed to identify small molecules that reduce access by endotoxin to the Asp299 residue, or modulate the kinetics of endotoxin binding in the region containing Asp299 in such a way that it more closely resembles the kinetics of endotoxin binding to baboon TLR4 protein. For example, mammals that are susceptible to LPS-mediated response (i.e., those with Asp299) could be administered an appropriate dose of the candidate agent to determine if it attenuates the sepsis symptoms normally associated with exposure to LPS.

Example 8. Screening antibody candidates

Antibodies (or modified antibodies or antibody fragments) are isolated/designed that bind to the extracellular region of human TLR4 such that access by bacterial LPS is diminished to Asp299. Such antibodies are directed to an epitope comprising Asp299. Preferably the epitope is 10 or less residues in length. Creation or isolation of such antibodies is understood by those skilled in the art. Uehori et al. (2003) Infect. Immun. 71(8):4238, describe antibodies to TLR4 which inhibited bacterial cell wall skeleton mediated NF-kB activation by 80%. Also see akashi et al. (2000) J. Immunol. 164:3471-75.

Claims:

1. A method of identifying a nucleotide change in a TLR4 polynucleotide sequence of an Old World monkey wherein said change may be associated with reduced sensitivity to Gram-negative bacterial infection, comprising the step of:

comparing the TLR4 polynucleotide sequence of the Old World monkey with corresponding TLR4 polynucleotide sequence of a human to identify a polynucleotide change in said Old World monkey's TLR4 sequence that is evolutionarily meaningful, whereby said evolutionarily meaningful change may be associated with reduced sensitivity to Gram-negative bacterial infection.

2. The method of claim 1 wherein the Old World monkey is selected from the group consisting of rhesus monkey and baboon.

3. The method of claim 2 wherein the evolutionarily meaningful change is from Asp299 in the human to Asn299 in the rhesus monkey or baboon.

4. The method of claim 3, wherein the evolutionarily meaningful change is associated with the reduced sensitivity to Gram-negative bacterial infection by the step comprising:

analyzing the functional effect of the Asp299Asn change in a model system.

5. The method of claim 4, wherein said model system is *in vivo*, *ex vivo* or *in vitro*.

6. A method of identifying a therapeutic agent that reduces sensitivity to Gram-negative bacterial infection, comprising:

(a) contacting candidate agents with human TLR4 polypeptide; and

(b) identifying a therapeutic agent that interacts with the TLR4 polypeptide to substantially reduce sensitivity to Gram-negative bacterial infection.

7. The method of claim 6 wherein said interaction with TLR4 polypeptide occurs at Asp299.

8. The method of claim 6, wherein said substantial reduction in sensitivity to Gram-negative bacterial infection is determined by an indicator selected from the group consisting of:

(a) elimination or substantial reduction in host systemic inflammatory response to LPS in a human, non-human primate, or suitable animal model; and

(b) elimination or reduced severity of central nervous system dysfunction, adult respiratory distress syndrome, liver failure, acute renal failure, and/or disseminated intravascular coagulation in a human, non-human primate, or suitable animal model.

9. A method for treating sepsis, severe sepsis or septic shock in a primate, comprising:

administering to a primate in need thereof an effective dose of a therapeutic agent identified according to the method of claim 6.

10. A method for treating asthma in a primate, comprising:

administering to a primate in need thereof an effective dose of a therapeutic agent identified according to the method of claim 6.

11. A therapeutic agent identified according to the method of claim 6.

12. A composition comprising a polynucleotide selected from the group consisting of chimpanzee *TLR4* polynucleotide (SEQ. ID. NO. __), gorilla *TLR4* polynucleotide (SEQ. ID. NO. __), gibbon *TLR4* polynucleotide (SEQ. ID. NO. __), rhesus monkey *TLR4* polynucleotide (SEQ. ID. NO. __), capuchin *TLR4* polynucleotide (SEQ. ID. NO. __), squirrel monkey *TLR4* polynucleotide (SEQ. ID NO. __), and baboon *TLR4* polynucleotide (SEQ. ID NO.).

13. A composition comprising a polypeptide selected from the group consisting of chimpanzee TLR4 polypeptide (SEQ. ID. NO. __), gorilla TLR4 polypeptide (SEQ. ID. NO. __), gibbon TLR4 polypeptide (SEQ. ID. NO. __), rhesus monkey TLR4 polypeptide (SEQ. ID. NO. __), capuchin TLR4 polypeptide (SEQ. ID. NO. __), squirrel monkey TLR4 polypeptide (SEQ. ID NO. __), and baboon TLR 4 polypeptide (SEQ. ID NO.).

Human	CNLTIEEFRLTYLD-YYLDdIIDLFNCLANASSFSL
Human null	CNLTIEEFRLTYLD-YYLDgIIDLFNCLANASSFSL
Chimpanzee	CNLTIEEFRLTYLD-YYLDdIIDLFNCLANASSFSL
Bonobo	CNLTIEEFRLTYLD-YYLDdIIDLFNCLANASSFSL
Gorilla	CNLTIEEFRLTYLD-YYLDdIIDLFNCLANASSFSL
Orangutan	CNLTIEEFRLAYLD-YYLDdIIDLFNCLANVSSFSL
Gibbon	CNLTIEEFRLTYLD-YYLDdIIDLFNCLANASSFSL
Baboon	CNLTIEEFRLTYLD-YYLDnIIDLFNCLANASSFSL
Rhesus	CNLTIEEFRLTYLD-YYLDnIIDLFNCLANASSFSL
Horse	HNLTIEEFRLAYIDNYSSKdSIDLLNCLADISKISL
Cow	CNLTIEQFRIAYLDKFSGDd-TDLFNCLANVSVISL
Cat	CNLIIEKFRIAYFDKFS-EdAIDSFNCLANVSTISL
Dog	CNLTIEKFRIAYFDSFS-KdTTNLFNQLVNISAISL
Hamster	CKVTIEEFRFTYANEFS-EdITD-FDCLANVSAMSL
Rat	CNVSIDEFRLTYINHFS-DdIYN-LNCLANISAMSF
Mouse	CDVTIDEFRLTHTNDFS-DdI-VKFHCLANVSAMSL

Figure 1.

Baboon CDS

GTGGTTCCTAACATTACTTATCAATGCATGGAGCTGAATTTCTACAAAATCCCCGACAACA
TCCCCTTCTCAACCAAGAACCTGGACCTGAGCTTTAATCCCCTGAGGCATTTAGGCAGCTA
TAGCTTCCTCCGTTTTCCAGAACTGCAGGTGCTGGATTTATCCAGGTGTGAAATCCAGACA
ATTGAAGATGGGGCATATCAGAGCCTAAGCCACCTCTCCACCTTAATATTGACAGGAAAC
CCCATCCAGAGTTTAGCCCTGGGAGCCTTTTCTGGACTATCAAGTTTACAGAAGCTGGTGG
CTGTGGAGACAAATCTAGCATCTCTAGAGAACTTCCCCATTGGACATCTCAAAACTTTGAA
AGAACTTAATGTGGCTCACAATCTTATCCAGTCTTTCAAATTACCTGAGTATTTTTCTAATC
TGACCAATCTAGAGCACTTGGACCTTTCCAGTAACAAGATTCAAAATATTTATTGCAAAGA
CTTGCAGGTTCTACATCAAATGCCCCTACCCAATCTCTCTTTAGACCTGTCCCTGAACCCTA
TAAACTTTATCCAACCAGGTGCATTTAAAGAAATTAGGCTTCATAAGCTGACTTTGAGAAG
TAATTTTGATGATTTAAATGTAATGAAAACTTGTATTCAAGGTCTGGCTGGTTTAGAAGTC
CATCGTTTGGTTCTGGGAGAATTTAGAAATGAAAGAAACTTGGAAGAGTTTGACAAATCT
GCTCTGGAGGGATTGTGCAATTTGACCATTGAAGAATTCCGATTAACATACTTAGACTACT
ACCTCGATAATATTATTGACTTATTTAATTGTTTGGCAAATGCTTCTTCATTTTCCCTGGTG
AGTGTGAATATTAAAAGGGTAGAAGACTTTTCTTATAATTTCAGATGGCAACATTTAGAAT
TAGTTAACTGTAAATTTGAACAGTTTCCCACATTGGAACTCGAATCTCTCAAAAGGCTTAC
TTTCACTGCCAACAAAGGTGGGAATGCCTTTTCAGAAGTTGATCTACCAAGCCTTGAGTTT
CTAGATCTCAGTAGAAATGGCTTGAGTTTCAAAGGTTGCTGTTCTCAAAGTGATTTTGGGA
CAACCAGCCTAAAGTATTTAGATCTGAGCTTCAATGATGTTATTACCATGGGTTCAAACTT
CTTGGGCTTAGAACAACTAGAACATCTGGATTTCCAGCATTCCAATTTGAAACAGATGAGT
CAATTTTCAGTATTCCTATCACTCAGAAACCTCATTTACCTTGACATTTCTCATACTCACAC
CACAGTTGCTTTCAATGGCATTTTCGATGGCTTGCTCAGTCTCAAAGTCTTAAAAATGGCT
GGCAATTCTTTCCAGGAAAACTTCCTTCCAGATATCTTCACAGATCTGAAAAACTTGACCT
TCCTGGACCTCTCTCAGTGTCAACTGGAGCAGTTGTCTCCAACAGCATTTGACACACTCAA
CAAGCTTCAGGTACTAAATATGAGCCACAACAACTTCTTTTCATTGGATGTGTTTCCTTAT
AAGTGTCTGCCCTCCCTCCAGGTTCTCGATTACAGTCTCAATCACATAATGACTTCCAAAA
ACCAGGAACCTCAGCATTTTCCAAGTAGTCTAGCTTTCTTAAATCTTACTCAGAATGACTT
TGCTTGTACTTGTGAACACCAGAGTTTCCTGCAGTGGATCAAGGACCAGAGGCAGCTCTTG
GTGGAAGCTGAACGAATGGAATGTGCAACACCTTCAGATAAACAGGGCATGCCTGTGCTG
AGTGTGAATATTACCTGTCAGATGAATAAGACCATCATTGGTGTGTCTGTGTTCAGTGTGC
TTGTGGTATCTGTTGTAGCAGTTCTGGTCTATAAGTTCTATTTTCACCTGATGCTTCTTGCT
GGCTGCATAAAGTATGGTAGAGGTGAAAACATCTATGATGCCTTTGTTATCTACTCAAGCC
AGGATGAGGACTGGGTAAGGAATGAGCTAGTAAAGAATTTAGAAGAAGGGGTGCCTCCC
TTTCAGCTCTGCCTTCACTACAGAGACTTTATTCCCGGTGTGGCCATTGCTGCAAACATCA
TCCATGAAGGTTTCCATAAAAGCCGAAAGGTGATTGTTGTGGTGTCCCAGCACTTCATCCA
GAGCCGCTGGTGTATCTTTGAATATGAGATTGCTCAGACCTGGCAGTTTCTGAGCAGTCGT
GCAGGCATAATCTTCATTGTCCTGCAGAAGGTGGAGAAGACCCTGCTCAGGCAGCAGGTG
GAGCTGTACCGCCTTCTCAGCAGGAACACTTACCTGGAGTGGGAGGACAGTGTCCTAGGG
CAGCACATCTTCTGGAGACGACTCAGAAAAGCCCTGTTGGATGGCAGATCGTGGAATCCA
GAAGAACAGTAG

FIGURE 2

Bonobo

GTGGTTCCTAATATTACTTATCAATGCATGGAGCTGAATTTCTACAAAATCCCCGACAACC
TCCCCTTCTCAACCAAGAACCTGGACCTGAGCTTTAATCCCCTGAGGCATTTAGGCAGCTA
TAGCTTCTTCAGTTTCCCAGAACTGCAGGTGCTGGATTTATCCAGGTGTGAAATCCAGACA
ATTGAAGATGGGGCATATCAGAGCCTAAGCCACCTCTCCACCTTAATATTGACAGGAAAC
CCCATCCAGAGTTTAGCCCTGGGAGCCTTTTCTGGACTATCAAGTTTACAGAAGCTGGTGG
CTGTGGAGACAAATCTAGCATCTCTAGAGAACTTCCCCATTGGACATCTCAAAACTTTGAA
AGAACTTAATGTGGCTCACAATCTTATCCAATCTTTCAAATTACCTGAGTATTTTTCTAATC
TGACCAATCTAGAGCACTTGGACCTTTCCAGCAACAAGATTCAAAGTATTTATTGCACAGA
CTTGCGGGTTCTACATCAAATGCCCCTACTCAATCTCTCTTTAGACCTGTCCCTGAACCCTA
TGAACTTTATCCAACCAGGTGCATTTAAAGAAATTAGGCTTCATAAGCTGACTTTGAGAAA
TAATTTTGATAGTTTAAATGTAATGAAAACTTGTATTCAAGGTCTGGCTGGTTTAGAAGTC
CATCGTTTGGTTCTGGGAGAATTTAGAAATGAAGAAAACTTGGAAAAGTTTGACAAATCT
GCTCTAGAGGGCCTGTGCAATTTGACCATTGAAGAATTCCGATTAGCATACTTAGACTACT
ACCTCGATGATATTATTGACTTATTTAATTGTTTGACAAATGTTTCTTCATTTTCCCTGGTG
AGTGTGACTATTAAAAGCGTAAAAGACTTTTCTTATAATTTCGGATGGCAACATTTAGAAT
TAGTTAAGTGTAAATTTGGACAGTTTCCCACATTGAAACTCAAATCTCTCAAAAGGCTTAC
TTTCACTTCCAACAAAGGTGGGAATGCTTTTTCAGAAGTTGATCTACCAAGCCTTGAGTTT
CTAGATCTCAGTAGAAATGGCTTGAGTTTCAAAGGTTGCTGTTCTCAAAGTGATTTTGGGA
CAACCAGCCTAAAGTATTTAGATCTGAGCTTCAATGGTGTTATTACCATGAGTTCAAACTT
CTTGGGCTTAGAACAACTAGAACATCTGGATTTCCAGCATTCCAATTTGAAACAAATGAGT
GAGTTTTCAGTATTCCTATCACTCAGAAACCTCATTTACCTTGACATTTCTCATACTCACAC
CAGAGTTGCTTTCAATGGCATCTTCAATGGCTTGTCCAGTCTCGAAGTCTTGAAAATGGCT
GGCAATTCTTTCCAGGAAAACTTCCTTCCAGATATCTTCACAGAGCTGAGAAACTTGACCT
TCCTGGACCTCTCTCAGTGTCAACTGGAGCAGTTGTCTCCAACAGCATTTAACTCACTCTC
CAGTCTTCAGGTACTAAATATGAGCCACAACAACTTCTTTTCATTGGATACGTTTCCTTAT
AAGTGTCTGAACTCCCTCCAGGTTCTTGATTACAGTCTCAATCACATAATGACTTCCAAAA
AACAGGAACTACAGCATTTTCCAAGTAGTCTAGCTTTCTTAAATCTTACTCAGAATGACTT
TGCTTGTACTTGTGAACACCAAAGTTTCCTGCAATGGATCAAGGACCAGAGGCAGCTCTTG
GTGGAAGTTGAACGAATGGAATGTGCAACACCTTCAGATAAGCAGGGCATGCCTGTGCTG
AGTTTGAATATCACCTGTCAGATGAATAAGACCATCATTGGTGTGTCGGTCCTCAGTGTGC
TTGTAGTATCTGTTGTAGCAGTTCTGGTCTATAAGTTCTATTTTCACCTGATGCTTCTTGCT
GGCTGCATAAAGTATGGTAGAGGTGAAAACATCTATGATGCCTTTGTTATCTACTCAAGCC
AGGATGAGGACTGGGTAAGGAATGAGCTAGTAAAGAATTTAGAAGAAGGGGTGCCTCCA
TTTCAGCTCTGCCTTCACTACAGAGACTTTATTCCCGGTGTGGCCATTGCTGCCAACATCAT
CCATGAAGGTTTCCATAAAAGCCGAAAGGTGATTGTTGTGGTGTCCCAGCACTTCATCCAG
AGCCGCTGGTGTATCTTTGAATATGAGATTGCTCAGACGTGGCAGTTTCTGAGCAGTCGTG
CTGGTATCATCTTCATTGTCCTGCAGAAGGTGGAGAAGACCCTGCTCAGGCGGCAGGTGG
AGCTGTACCGCCTTCTYAGCAGGAACACTTACCTGGAGTGGGAGGACAGTGTCCTGGGGC
GGCACATCTTCTGGAGACGACTCAGAAAAGCCCTGCTGGATGGTAAATCATGGAATCCAG
AAGGAACAGTGGGTACAGGATGCAATTGGCAGGAAGCAACATCTATCTGA

FIGURE 3

Gibbon

GTGGTTCCTAACATTACTTATCAATGCATGGAGCTGAATTTCTACAAAATCCCCGACAACC
TCCCCTTCTCAACCAAGAACCTGGACCTGAGCTTTAATCCCCTGAGGCATTTAGGCAGCTA
TAGCTTCTTCAGTTTCCCAGAACTGCAGGTGCTGGATTTATCCAGGTGTGAAATCCAGACA
ATTGAAGATGGGGCATATCAGAGCCTAAGCCTCCTCTCCACCTTAATATTGACAGGAAAC
CCCATCCAGAGTTTAGCTCTGGGAGCCTTTTCTGGACTATCAAGTTTACAGAAGCTAGTGG
CTGTGGAGACAAATCTAGCATCTCTAGAGAACTTCCCCATTGGACATCTCAAAACTTTGAA
AGAACTTAATGTGGCTCACAATCTTATCCAATCTTTCAAATTACCTGAGTATTTTTCTAATC
TGACCAATCTAGAGCACTTGGACCTTTCCAGCAACAAGATTCAAAGTATTTATTGCAAAG
ACTTGCAGGTTCTACATCAAATGCCCCTACTCAATCTCTCTTTAGACCTGTCCCTGAACCCT
ATGAACTTTATCCAACCAGGTGCATTTAAAGAAATTAGCCTTCRTAAGCTGACTTTAAGAA
ATAATTTTGATAGTTTAAATGTAATGAAAACTTGTATTCAAGGTCTGGCTGGTTTAGAAGT
CCATCGTTTGGTTCTGGGAGAATTTAGAAATGAAGGAAACTTGGAAGAGTTTGACAAATC
TGCTCTAGAGGGCCTGTGCAATTTGACCATTGAAGAATTCCGATTAGCATACTTAGACCAC
TACCTCGATGATATTATTGACTTATTTAATTGTTTGGCAAATGTTTCTTCATTTTCCCTGGT
GAGTGTGACTATTAAAAGGGTAGAAGACTTTTCTTATAATTTCGGATGGCAACATTTAGAA
TTAGTTAACTGTAAATTTGGACAGTTTCCCACATTGAACCTCAAATCTCTCAAAAGGCTTA
CTTTCACTGCCAACAGAGGTGGGAATGCTTTTTCAGAAGTTGATCTACCAAGCCTTGAGTT
TCTAGATCTCAGTAGAAATGGCTTGAGTTTCAAAGGTTGCTGTTCTCAAAGTGATTTTGGG
ACAAACAGCCTAAAGTATTTAGATCTGAGCTTCAATGATGTTATTACCATGAGTTCAAACT
TCTTGGGCTTAGAACAGCTAGAACATCTGGATTTGCAGCATTCCAATTTGAAACAAATGA
GTGAATTTTCAGTATTCCTATCACTCAGAAACCTCATTTACCTTGACATTTCTCATACTCAC
ACCAGAGTTGCTTTCAATGGCATCTTCAATGGCTTGTCCAATCTCGAAGTCTTGAAAATGG
CTGGCAATTCTTTCCAGGAAAACTTCCTTCCAGATATCTTCACAGAGCTGAGAAACTTGAC
CTTCCTGGACCTCTCTCAGTGTCAACTGGAGCAATTGTCTCCAACAGCATTTAACTCACTC
TCCAGTCTTCAGGTACTAAATATGAGCCACAACAACTTCTTTTCATTGGATACGTTTCCTTA
TAAGTGTCTGAACTCCCTCCAGGTTCTTGATTACAGTCTCAATCACATAATGACTTCCAAA
AAACAGGAACTACAGCGTTTTCCAAGTAGTCTAGCCTTCTTAAATCTTACTCAGAATGACT
TTGCTTGTACTTGTGAACACGAGAGTTTCCTGCAGTGGATCAAGGACCAGAGGCAGCTCTT
GGTGGAAGTTGAACGAATGGAATGTGCAACACCTTCAGATAAGCAGGGCATGCCTGTGCT
GAGTTTGAATATCACCTGTCAGATGAATAAGACCATCATTGGTGTGTCAGTCCTCAGTGTG
CTTGTAGTATCTGTTGTAGCAGTTCTGGTCTATAAGTTCTATTTTCACCTGATGCTTCTTGC
TGGCTGCATGAAGTATGGTAGAGGTGAAAACACCTATGATGCCTTTGTTATCTACTCCAGC
CAGGATGAGGACTGGGTAAGGAATGAGCTAGTAAAGAATTTAGAAGAAGGGGTGCCTCC
CTTTCAGCTCTGCCTTCACTACAGAGACTTTATTCCYGGTGTGGCCATTGCTGCCAACATC
ATCCATGAAGGTTTCCATAAAAGCCGAAAGGTGATTGTTGTGGTGTCCCAGCACTTCATCC
AGAGCCGCTGGTGTATCTTTGAGTATGAGATTGCTCAGACCTGGCAGTTTCTGAGCAGTCA
TGCTGGGATCATCTTCATTGTCCTGCAGAAGGTGGAGAAGACCCTGCTCAGGCAGCAGGT
GGAGCTGTACCGCCTTCTCAGCAGGAACACTTACCTGGAGTGGGAGGATAGTGTCCTGGG
GCGGCACATTTTCTGGAGACGACTCAGAAAAGCCCTGCTGGATGGTAAATCATGGAATCC
AGAAGGAACAGTGGGTACAGGATGCAATTAG

FIGURE 4

Gorilla

GTGGTTCCTAATATTACTTATCAATGCATGGAGCTGAATTTCTACAAAATCCCCGACAACC
TCCCCTTCTCAACCAAGAACCTGGACCTGAGCTTTAATCCCCTGAGGCATTTAGGCAGCTA
TAGCTTCTTCAGTTTCCCAGAACTGCAGGTGCTGGATTTATCCAGGTGTGAAATCCAGACA
ATTGAAGATGGGGCATATCAGAGCCTAAGCCACCTCTCCACCTTAATATTGACAGGAAAC
CCCATCCAGAGTTTAGCCCTGGGAGCCTTTTCTGGACTATCAAGTTTACAGAAGCTGGTGG
CTGTGGAGACAAATCTAGCATCTCTAGAGAACTTCCCCATTGGACATCTCAAAACTTTGAA
AGAACTTAATGTGGCTCACAATCTTATTCAATCTTTCAAATTACCTGAGTATTTTTCTAATC
TGACCAATCTAGAGTACTTGGACCTTTCCAGCAACAAGATTCAAAGTATTTATTGCACAGA
CTTGCGGGTTCTACATCAAATGCCCCTACTCAATCTCTCTTTAGACCTGTCCCTGAACCCTA
TGACCTTTATCCAACCAGGTGCATTTAAAGAAATTAGGCTTCATAAGCTGACTTTGAGAAA
TAATTTTGATAGTTTAAATGTAATGAAAACTTGTATTCAAGGTCTGGCTGGTTTAGAAGTC
CGTCGTTTGGTTCTGGGAGAATTTAGAAATGAAGGAAACTTGGAAAAGTTTGACAAATCT
GCTCTAGAGGGCCTGTGCAATTTGACCATTGAAGAATTCCGATTAGCATACTTAGACTACT
ACCTCGATGATATTATTGACTTATTTAATTGTTTGACAAATGTTTCTTCATTTTCCCTGGTG
AGTGTGACTATTGAAAGGGTAAAAGACTTTTCTTATAATTTCGGATGGCAACATTTAGAAT
TAGTTAACTGTAAATTTGGACAGTTTCCCACATTGAAACTCAAATCTCTCAAAAGGCTTAC
TTTCACTTCCAACAAAGGTGGGAATGCTTTTTCGGAAGTTGATCTACCAAGCCTTGAGTTT
CTAGATCTCAGTAGAAATGGCTTGAGTTTCAAAGGTTGCTGTTCTCAAAGTGATTTTGGGA
CAACCAGCCTAAAGTATTTAGATCTGAGCTTCAATGGTGTTATTACCATGAGTTCAAACTT
CTTGGGCTTAGAACAACTAGAACATCTGGATTTCCAGCATTCCAATTTGAAACAAATGAGT
GAGTTTTCAGTATTCCTATCACTCAGAAACCTCATTTACCTTGACATTTCTCATACTCACAC
CAGAGTTGCTTTCAATGGCATCTTCAATGGCTTGTCCAGTCTCGAAGTCTTGAAAATGGCT
GGCAATTCTTTCCAGGAAAACTTCCTTCCAGATATCTTCACAGAGCTGAGAAACTTGACCT
TCCTGGACCTCTCTCAGTGTCAACTGGAGCAGTTGTCTCCAACAGCATTTAACTCACTCTC
CAGTCTTCAGGTACTAAATATGAGCCACAACAACTTCTTTTCATTGGATACGTTTCCTTAT
AAGTGTCTGAACTCCCTCCGGGTTCTTGATTACAGTCTCAATCACATAATGACTTCCAAAA
AACAGGAACTACAGCATTTTCCAAGCAGTCTAGCTTTCTTAAATCTTACTCAGAATGACTT
TGCTTGTACTTGTGAACACCAGAGTTTCCTGCAATGGATCAAGGACCAGAGGCAGCTCTTG
GTGGAAGTTGAACGAATGGAATGTGCAACACCTTCAGATAAGCAGGGCATGCCTGTGCTG
AGTTTGAATATCACCTGTCAGATGAATAAGACCATCATTGGTGTGTCGGTCCTCAGTGTGC
TTGTAGTATCTGTTGTAGCAGTTCTGGTCTATAAGTTCTATTTTCACCTGATGCTTCTTGCT
GGCTGCATAAAGTATGGTAGAGGTGAAAACGTCTATGATGCCTTTGTTATCTACTCAAGCC
AGGATGAGGACTGGGTAAGGAATGAGCTAGTAAAGAATTTAGAAGAAGGGGTGCCTCCA
TTTCAGCTCTGCCTTCACTACAGAGACTTTATTCCCGGTGTGGCCATTGCTGCCAACATCAT
CCATGAAGGTTTCCATAAAAGTCGAAAGGTGATTGTTGTGGTGTCCCAGCACTTCATCCAG
AGCCGCTGGTGTATCTTTGAATATGAGATTGCTCAGACCTGGCAGTTTCTGAGCAGTCGTG
CTGGTATCATCTTCATTGTCCTGCAGAAGGTGGAGAAGACCCTGCTCAGGCAGCAGGTGG
AGCTGTACCGCCTTCTCAGCAGGAACACTTACCTGGAGTGGGAGGACAGTGTCCTGGGGC
GGCACATCTTCTGGAGACGACTCAGAAAAGCCCTGCTGGATGGTAAATCATGGAATCCAG
AAGGAACAGTGGGTACAGGATGCAATTGGCAGGAAGCAACATCTATCTGA

FIGURE 5

Rhesus monkey

GTGGTTCCTAATATTACTTATCAATGCATGGAGCTGAATTTCTACAAAATCCCCGACAACC
TCCCCTTCTCAACCAAGAACCTGGACCTGAGCTTTAATCCCCTGAGGCATTTAGGCAGCTA
TAGCTTCTTCAGTTTCCCAGAACTGCAGGTGCTGGATTTATCCAGGTGTGAAATCCAGACA
ATTGAAGATGGGGCATATCAGAGCCTAAGCCACCTCTCCACTTTAATATTGACAGGAAAC
CCCATCCAGAGTTTAGCCCTGGGAGCCTTTTCTGGACTATCAAGTTTACAGAAGCTGGTGG
CTGTGGAGACAAATCTAGCATCTCTAGAGAACTTCCCCATTGGACATCTCAAAACTTTGAA
AGAACTTAATGTGGCTCACAATCTTATCCAGTCTTTCAAATTACCTGAGTATTTTTCTAATC
TGACCAATCTAGAGCACTTGGACCTTTCCAGTAACAAGATTCAAAATATTTATTGCAAAGA
CTTGCAGGTTCTACATCAAATGCCCCTATCCAATCTCTCTTTAGACCTGTCCCTGAACCCTA
TAAACTTTATCCAACCAGGTGCATTTAAAGAAATTAGGCTTCATAAGCTGACTTTGAGAAG
TAATTTTGATGATTTAAATGTAATGAAAACTTGTATTCAAGGTCTGGCTGGTTTAGAAGTC
CATCGTTTGGTTCTGGGAGAATTTAGAAATGAAAGAAACTTGGAAGAGTTTGACAAATCT
TCTCTGGAGGGATTGTGCAATTTGACCATTGAAGAATTCCGATTAACATACTTAGACTACT
ACCTCGATAATATTATTGACTTATTTAATTGTTTGGCAAATGTTTCTTCATTTTCCCTGGTG
AGTGTGAGTATTAAAAGGGTAGAAGACTTTTCTTATAATTTCAGATGGCAACATTTAGAAT
TAGTTAACTGTAAATTTGAACAGTTTCCCACATTGGAACTCGAATCTCTCAAAAGGCTTAC
TTTCACTGCCAACAAAGGTGGGAATGCTTTTTCAGAAGTTGATCTACCAAGCCTTGAGTTT
CTAGATCTCAGTAGAAATGGCTTGAGTTTCAAAGGTTGCTGTTCTCAAAGTGATTTTGGGA
CAACCAGCCTAAAGTATTTAGATCTGAGCTTCAATGATGTTATTACCATGAGTTCAAACTT
CTTGGGCTTAGAAAAACTAGAACATCTGGATTTCCAGCATTCCAATTTGAAACAGATGAG
TCAATTTTCAGTATTCCTATCACTCAGAAACCTCATTTACCTTGACATTTCTCATACTCACA
CCAGAGTTGCTTTCAATGGCATCTTCGATGGCTTGCTCAGTCTCAAAGTCTTAAAAATGGC
TGGCAATTCTTTCCAGGAAAACTTCCTTCCAGATATCTTCACAGATCTGAAAAACTTGACC
TTCCTGGACCTCTCTCAGTGTCAATTGGAGCAGTTGTCTCCAACAGCATTTGACACACTCA
ACAAGCTTCAGGTACTAAATATGAGCCACAACAACTTCTTTTCATTGGATACGTTTCCTTA
TAAGTGTCTGCCCTCCCTCCAGGTTCTCGATTACAGTCTCAATCACATAATGACTTCCAAC
AACCAGGAACTACAGCATTTTCCAAGTAGTCTAGCTTTCTTAAATCTTACTCAGAATGACT
TTGCTTGTACTTGTGAACACCAGAGTTTCCTGCAGTGGATCAAGGACCAGAGGCAGCTCTT
GGTGGAAGCTGAACGAATGGAATGTGCAACACCTTCAGATAAACAGGGCATGCCGGTGCT
GAGTTTGAATATTACCTGTCAGATGAATAAGACCATCATTGGTGTGTCTGTGTTCAGTGTG
CTTGTGGTATCTGTTGTAGCAGTTCTGGTCTATAAGTTCTATTTTCACCTGATGCTTCTTGC
TGGCTGCATAAASTATGGTAGAGGTGAAAACATCTATGATGCCTTTGTTATCTACTCAAGC
CAGGATGAGGACTGGGTAAGGAATGAACTAGTAAAGAATTTAGAAGAAGGGGTGCCTCC
CTTTCAGCTCTGCCTTCACTACAGAGACTTTATTCCCGGTGTGGCCATTGCTGCAAACATC
ATCCATGAAGGTTTCCATAAAAGCCGAAAGGTGATTGTTGTGGTGTCCCAGCACTTCATCC
AGAGCCGCTGGTGTATCTTTGAATATGAGATTGCTCAGACCTGGCAGTTTCTGAGCAGTCG
TGCAGGCATAATCTTCATTGTCCTGCAGAAGGTGGAGAAGACCCTGCTCAGGCAGCAGGT
GGAGCTGTACCGCCTTCTCAGCAGGAACACTTACCTGGAGTGGGAGGACAGTGTCCTGGG
GCAGCACATCTTCTGGAGACGACTCAGAAAAGCCCTGTTGGATGGCAGATCGTGGAATCC
AGAAGAACAGTAG

FIGURE 6

Chimpanzee

GTGGTTCCTAATATTACTTATCAATGCATGGAGCTGAATTTCTACAAAATCCCCGACAACC
TCCCCTTCTCAACCAAGAACCTGGACCTGAGCTTTAATCCCCTGAGGCATTTAGGCAGCTA
TAGCTTCTTCAGTTTCCCAGAACTGCAGGTGCTGGATTTATCCAGGTGTGAAATCCAGACA
ATTGAAGATGGGGCATATCAGAGCCTAAGCCACCTCTCCACCTTAATATTGACAGGAAAC
CCCATCCAGAGTTTAGCCCTGGGAGCCTTTTCTGGACTATCAAGTTTACAGAAGCTGGTGG
CTGTGGAGACAAATCTAGCATCTCTAGAGAACTTCCCCATTGGACATCTCAAAACTTTGAA
AGAACTTAATGTGGCTCACAATCTTATCCAATCTTTCAAATTACCTGAGTATTTTTCTAATC
TGACCAATCTAGAGCACTTGGACCTTTCCAGCAACAAGATTCAAAGTATTTATTGCACAGA
CTTGCGGGTTCTACATCAAATGCCCCTACTCAATCTCTCTTTAGACCTGTCCCTGAACCCTA
TGAACTTTATCCAACCAGGTGCATTTAAAGAAATTAGGCTTCATAAGCTGACTTTGAGAAA
TAATTTTGATAGTTTAAATGTAATGAAAACTTGTATTCAAGGTCTGGCTGGTTTAGAAGTC
CATCGTTTGGTTCTGGGAGAATTTAGAAATGAAGGAAACTTGGAAAAGTTTGACAAATCT
GCTCTAGAGGGCCTGTGCAATTTGACCATTGAAGAATTCCGATTAGCATACTTAGACTACT
ACCTCGATGATATTATTGACTTATTTAATTGTTTGACAAATGTTTCTTCATTTTCCCTGGTG
AGTGTGACTATTAAAAGCGTAAAAGACTTTTCTTATAATTTCGGATGGCAACATTTAGAAT
TAGTTAACTGTAAATTTGGACAGTTTCCCACATTGAAACTCAAATCTCTCAAAAGGCTTAC
TTTCACTTCCAACAAAGGTGGGAATGCTTTTTCAGAAGTTGATCTACCAAGCCTTGAGTTT
CTAGATCTCAGTAGAAATGGCTTGAGTTTCAAAGGTTGCTGTTCTCAAAGTGATTTTGGGA
CAACCAGCCTAAAGTATTTAGATCTGAGCTTCAATGGTGTTATTACCATGAGTTCAAACTT
CTTGGGCTTAGAACAACTAGAACATCTGGATTTCCAGCATTCCAATTTGAAACAAATGAGT
GAGTTTTCAGTATTCCTATCACTCAGAAACCTCATTTACCTTGACATTTCTCATACTCACAC
CAGAGTTGCTTTCAATGGCATCTTCAATGGCTTGTCCAGTCTCGAAGTCTTGAAAATGGCT
GGCAATTCTTTCCAGGAAAACTTCCTTCCAGATATCTTCACAGAGCTGAGAAACTTGACCT
TCCTGGACCTCTCTCAGTGTCAACTGGAGCAGTTGTCTCCAACAGCATTTAACTCACTCTC
CAGTCTTCAGGTACTAAATATGAGCCACAACAACTTCTTTTCATTGGATACGTTTCCTTAT
AAGTGTCTGAACTCCCTCCAGGTTCTTGATTACAGTCTCAATCACATAATGACTTCCAAAA
AACAGGAACTACAGCATTTTCCAAGTAGTCTAGCTTTCTTAAATCTTACTCAGAATGACTT
TGCTTGTACTTGTGAACACCAAAGTTTCCTGCAATGGATCAAGGACCAGAGGCAGCTCTTG
GTGGAAGTTGAACGAATGGAATGTGCAACACCTTCAGATAAGCAGGGCATGCCTGTGCTG
AGTTTGAATATCACCTGTCAGATGAATAAGACCATCATTGGTGTGTCGGTCCTCAGTGTGC
TTGTAGTATCTGTTGTAGCAGTTCTGGTCTATAAGTTCTATTTTCACCTGATGCTTCTTGCT
GGCTGCATAAAGTATGGTAGAGGTGAAAACATCTATGATGCCTTTGTTATCTACTCAAGCC
AGGATGAGGACTGGGTAAGGAATGAGCTAGTAAAGAATTTAGAAGAAGGGGTGCCTCCA
TTTCAGCTCTGCCTTCACTACAGAGACTTTATTCCCGGTGTGGCCATTGCTGCCAACATCAT
CCATGAAGGTTTCCATAAAAGCCGAAAGGTGATTGTTGTGGTGTCCCAGCACTTCATCCAG
AGCCGCTGGTGTATCTTTGAATATGAGATTGCTCAGACCTGGCAGTTTCTGAGCAGTCGTG
CTGGTATCATCTTCATTGTCCTGCAGAAGGTGGAGAAGACCCTGCTCAGGCGGCAGGTGG
AGCTGTACCGCCTTCTCAGCAGGAACACTTACCTGGAGTGGGAGGACAGTGTCCTGGGGC
GGCACATCTTCTGGAGACGACTCAGAAAAGCCCTGCTGGATGGTAAATCATGGAATCCAG
AAGGAACAGTGGGTACAGGATGCAATTGGCAGGAAGCAACATCTATCTGA

FIGURE 7

Capuchin

TGTGAAATCCACACAATTGAAGATGGTGCATATCAGAGCCTAAGCCACCTCTCCACCTTA
ATATTGACAGGAAATCCTATCCAGAATTTAGCCCTGGGAGCCTTTTCTGGACTATCAAGTT
TACAGAAACTGGTAGCTGTGGAGACACATCTGTTATCGCTAGAAAGCTTCCCCATTGGAC
ATCTCAAAACTTTGAAGGACCTTAATGTGGCTCACAATCTAATCCAATCTTTCAAATTACC
TGAGTATTTTTCTAATCTGACCAATCTAGAGCACTTGGACCTTTCTAGTAACAATATTCAA
AATATTTATTGCAAAGACTTGCAGGTTCTACATCAAATGCCCCTACTCAATCTCTCTTTAG
ACCTGTCCCTGAACCCTATAAACTTTATTCAGCCAGGTGCATTTAAAGAAATTAGGCTCCG
TAAGCTGACTTTGAGAAATAATTTTGATAGTTTAAATGTAATGAAAACTTGCATTCACGGT
CTGGCTGGGTTAGAAGTCCATCGTTTGGTTCTGGGAGAATTTAGAAATGAAAGAAATATT
GAAGACTTTGACAAATCTGCTCTGGAGGGCCTGTGCAATTTGACCATCAAAGAATTCCGA
TTAGCATACTTAGACAACTTTCCAGATGATATTATTGACTTATTTAATTGTTTGGTAAATGT
TTCTTCATTTTCCCTGTTGAGTGTGTATATTAAAAGAGTAGAAGACTTTTCTTATAATTTCA
GATGGCAACATTTAGAATTAGTTAACTGTATATTTCAACAGTTTCCTCCACTGAAACTCAA
ATCTCTCAAAAGGCTTACTTTCAGTAAAAACAAAGGTAGGAATCATTTTGCAGAAGTTGA
TCTGCCAAGCCTTGAGTTTCTAGATCTCAGTAGAAATGGCTTGAGTTTCAAAGGTTGCTGT
TCTCAATCTGATTTTGGGACGACCAGCCTAAAGTATTTAGATCTGAGCTTCAATGATGTTA
TTACCATGAGTTCAAACTTCTTAGGCTTAGAACAACTAGAACACTTGGATTTCCAGCATTC
CAATTTGAAACAAATGAGTGAGTTTTCAGTATTTCTATCACTCAGAAACCTCATTTACCTT
GACATTTCTCATACTCACACCAGAGTTGCTTTCAATGGCATCTTTAATGGCTTGTTCAGTCT
CAAAGTCTTGAAAATGGCTGGAAATTCTTTCCAGCAAAACTTCCTTGCAGATATCTTCACA
GATCTGAATAACTTGATATTCCTGGACCTTTCTGAGTGTCAACTGGAGCAGTTGTCTCCAA
CAGCATTTGACTCACTTCCCAGACTTCAGATACTAAATATGAGCCACAACAAGTTCTTTGC
ATTGGATACATTTCCTTATAAGCATCTCTACTCCCTCCACGTTCTGGATTACAGTCTCAATC
ACATAGGGACTTCCAAAAATCAGGAACTACAGCATTTTCCAAGTAGTCTAGCTTTCTTAAA
TCTTACTCAAAATGACTTTGCTTGTACTTGTGAACACCAGAGTTTCCTGCAGTGGATCAAG
GACCAGAGGCGGCTATTGGTGGAAGTTGAACGAATGGAATGCGCAACACCTTTAAATAGG
AAGGGCATACCTGTGCTGAGTTTGAATATCACCTGTCAGATGAGTAAGACCATCATTGGT
GTGTCAGTGCTCAGTGTGCTTGTGGTATCTGTTGTAGCAGTTCTGGTCTATAAGTTCTATTT
TCACCTGATGCTTCTTGCTGGCTGCATAAAGTATGGTAGAGGTGAAAACACCTATGATGCC
TTTGTTATCTACTCAAGCCAGGATGAGGACTGGGTAAGGAATGAACTAGTAAAGAATTTA
GAAGAAGGGGTGCCTCCTTTTCAGCTCTGCCTTCACTACAGAGACTTTATTCCCGGTGTGG
CCATTGCTGCCAACATCATCCATGAAGGTTTCCATAAAAGCCGAAAGGTGATTGTTGTGGT
ATCCCAGCACTTCATCCAGAGCCGCTGGTGTATCTTTGAATATGAGATTGCTCAGACCTGG
CAGTTTCTGAGCAGTCGTGCTGGTATCATCTTCATTGTCCTGCAGAAGGTGGAGAAGTCCC
TGCTCAGGCAGCAGGTGGAGCTGTACCGCCTTCTCAGCAGGAACACCTACCTGGAGTGGG
AGGACAGTGTCCTGGGGAGGCATATCTTCTGGAGGCGACTCAGAAAAGCCCTGCTGAATG
GTAGACCGTGGAGTCCAGAAGGAACAGTGGGTGCAGGATGCGATTAG

FIGURE 8

Squirrel monkey

GTGGTTCCTAACGTTACTTATCAATGCATGGAACTGAATYTCTACAAAATCCCCGACAACA
TCCCCTTCTCAACTAAGAACCTGGACCTGAGCTTTAACCCCCTGAGGCATTTAGGCAGCCA
TAGCTTCTTCAATTTCCCAGAACTGCAGGTGCTGGATTTATCCAGGTGTGACATCCAGACA
ATCGAAGATGGGGCATATCAGAGCCTAAGCCACCTCTCCACCTTAATATTGACAGGAAAT
CCTATCCAGAATTTAGCCCTGGGAGCCTTTTCTGGACTATCAAGTTTACAGAAGCTGGTGG
CTGTGGAGACACATCTGTTATCACTAGAGAACTTCCCCATTGGACATCTCAAAACTTTGAA
GGACCTTAATGTGGCTCACAATCTAATCCAATCTTTCAAATTACCTGAGTATTTTTCTAATC
TGACCAATCTAGAGCACTTGGACCTTTCTAGTAACAATATTCAAAATATTTATTGCAAAGA
CTTGCAGGTTCTACATCAAATGCCCCTACTCAATCTCTCTTTAGACCTGTCCCTGAACCCTA
TAAACTTTATTCAACCAGGTGCGTTTAAAGAAATTAGGCTCCATAAGCTGACTTTGAGAAA
TAATTTTGATAGTTTAAATGCAATGAAAACTTGCATTCAAGGTCTGGCTGGGTTAGAAGTC
CATCGTTTGGTTCTGGGAGAATTTAGAAATGAAAGAAATATTGAAGACTTTGACAAATCT
GCTCTGGAGGGCCTGTGCAATTTGACCATTAATGAATTCCGATTAGCTTACTTAGATGACT
TTCTAGATGATATTATTGACTTATTTAACTGTTTAGCAAATGTTTCTTCATTTTCCCTGGTG
AATGTGCATATTAAAAGAGTAGAAGACTTTTCTTATAATTTTAGATGGCAACATTTAGAAT
TAGTTAACTGTGTATTTCAACAGTTTCCTCCACTGAAACTCAAATCTCTCAAAAGGCTTAC
TTTCACTGCCAACAAAGGTAGGAATCATTTTTCAGAAGTTGATCTTCCAAGCCTTGAGTTT
CTAGATCTCAGTAGAAATGGCTTGAGTTTCAAAGGTTGCTGTTCTCAATCTGATTTTGGGA
CGACCAGCCTAAAGTATTTAGATCTGAGCTTCAATGACGTTATTACCATGGGTTCAAACTT
CTTAGGCTTAGAACAACTAGAACACTTGGATTTCCAGCATTCCAATTTGAAACAAATGAGT
GAGTTTTCAGTATTCCTATCACTCAGAAACCTCATTTACCTTGACATTTCTCATACTCACAC
CAGAGTTGCTTTCAATGGCATCTTTAATGGCTTGTTCAGTCTCAAAGTCTTGAAAATGGCT
GGAAATTCTTTCCAGCAAAACTTCCTTGAAGATATCTTCACRGATCTGAATAACTTGATAT
TCCTGGACCTCTCTGAGTGTCAGCTGGAGCAGTTGTCTCCAACAGCATTTGACTCACTTCC
CAGACTTCGGATACTAAATATGAGCCACAACAACTTCTTTGCATTGGATACATTCCCTTAC
AAGCATCTCTACTCCCTCCAGGTTCTGGATTACAGTCTCAATCATATAGGGACTTCCAAAA
ATCAGGAACTGCAGCATTTTCCAAGTAGTCTAGCTTTCTTAAATCTTACTCAAAATGACTT
TGCTTGTACTTGTGAACACCAGAGTTTCCTGCAGTGGATCAAGGACCAGAGGCGGCTGTT
GGTGGAAGTTGAACAAATGGAATGTGCAACACCTTTAAATAGGAAGGGCATACCTGTGCT
GAGTTTGAATATCACCTGTCAGATGAGTAAGACTATCATTGGTGTGTCAGTGCTCAGTGTG
CTTGTGGTATCTGTTGTAGCAGTTCTGGTCTATAAGTTCTATTTTCACCTGATGCTTCTTGC
TGGCTGCATAAAGTATGGTAGAGGTGAAAACACCTATGATGCCTTTGTTATCTACTCAAGC
CAGGATGAGGACTGGGTAAGGAATGAACTAGTAAAGAATTTAGAAGAAGGGGTGCCTCC
CTTTCAGCTCTGCCTTCACTACAGAGACTTTATTCCCGGTGTGGCCATTGCTGCCAACATC
ATCCATGAAGGTTTCCATAAAAGCCGAAAGGTGATTGTTGTGGTATCTCAGCACTTCATCC
AGAGCCGCTGGTGTATCTTTGAATATGAGATTGCTCAGACCTGGCAGTTTCTGAGCAGTCG
TGCTGGTATCATCTTCATTGTCCTGCAGAAGGTGGAGAAGTCCCTGCTCAGGCAGCAGGTG
GAGCTGTACCGCCTTCTCAGCAGGAACACTTACCTGGAGTGGGAGGACAGTGTCCTGGGG
AGGCACATCTTCTGGAGACGACTCAGAAAAGCCCTGCTGGATGGTAGACCGTGGAATCCA
GAAGGAACAGTGGGTGCAGGATGCGAATAG

FIGURE 9

SEQUENCE LISTING

<110> Evolutionary Genomics LLC

<120> Development of Therapeutics for the Treatment of Endotoxin-Mediated Diseases

<130> GENO200.3.1

<150> 10/100,422
<151> 2002-03-18

<160> 40

<170> PatentIn version 3.2

<210> 1
<211> 2427
<212> DNA
<213> Pan troglodytes

<400> 1
gtggttccta atattactta tcaatgcatg gagctgaatt tctacaaaat ccccgacaac 60

ctccccttct caaccaagaa cctggacctg agctttaatc ccctgaggca tttaggcagc 120

tatagcttct tcagtttccc agaactgcag gtgctggatt tatccaggtg tgaaatccag 180

acaattgaag atggggcata tcagagccta agccacctct ccaccttaat attgacagga 240

aaccccatcc agagtttagc cctgggagcc ttttctggac tatcaagttt acagaagctg 300

gtggctgtgg agacaaatct agcatctcta gagaacttcc ccattggaca tctcaaaact 360

ttgaaagaac ttaatgtggc tcacaatctt atccaatctt tcaaattacc tgagtatttt 420

tctaatctga ccaatctaga gcacttggac ctttccagca acaagattca aagtatttat 480

tgcacagact tgcgggttct acatcaaatg cccctactca atctctcttt agacctgtcc 540

ctgaacccta tgaactttat ccaaccaggt gcatttaaag aaattaggct tcataagctg 600

actttgagaa ataattttga tagtttaaat gtaatgaaaa cttgtattca aggtctggct 660

ggtttagaag tccatcgttt ggttctggga gaatttagaa atgaaggaaa cttggaaaag 720

tttgacaaat ctgctctaga gggcctgtgc aatttgacca ttgaagaatt ccgattagca 780

tacttagact actacctcga tgatattatt gacttattta attgtttgac aaatgtttct 840

tcattttccc tggtgagtgt gactattaaa agcgtaaaag acttttctta taatttcgga 900

tggcaacatt tagaattagt taactgtaaa tttggacagt ttcccacatt gaaactcaaa 960

tctctcaaaa ggcttacttt cacttccaac aaaggtggga atgctttttc agaagttgat 1020

ctaccaagcc ttgagtttct agatctcagt agaaatggct tgagtttcaa aggttgctgt 1080

tctcaaagtg attttgggac aaccagccta aagtatttag atctgagctt caatggtgtt 1140

attaccatga gttcaaactt cttgggctta gaacaactag aacatctgga tttccagcat 1200

tccaatttga aacaaatgag tgagttttca gtattcctat cactcagaaa cctcatttac 1260

cttgacattt ctcatactca caccagagtt gctttcaatg gcatcttcaa tggcttgtcc 1320

agtctcgaag tcttgaaaat ggctggcaat tctttccagg aaaacttcct tccagatatc 1380

ttcacagagc tgagaaactt gaccttcctg gacctctctc agtgtcaact ggagcagttg 1440

tctccaacag catttaactc actctccagt cttcaggtac taaatatgag ccacaacaac 1500

ttcttttcat tggatacgtt tccttataag tgtctgaact ccctccaggt tcttgattac 1560

agtctcaatc acataatgac ttccaaaaaa caggaactac agcattttcc aagtagtcta 1620

gctttcttaa atcttactca gaatgacttt gcttgtactt gtgaacacca aagtttcctg 1680

caatggatca aggaccagag gcagctcttg gtggaagttg aacgaatgga atgtgcaaca 1740

ccttcagata agcagggcat gcctgtgctg agtttgaata tcacctgtca gatgaataag 1800

accatcattg gtgtgtcggt cctcagtgtg cttgtagtat ctgttgtagc agttctggtc 1860

tataagttct attttcacct gatgcttctt gctggctgca taaagtatgg tagaggtgaa 1920

aacatctatg atgcctttgt tatctactca agccaggatg aggactgggt aaggaatgag 1980

ctagtaaaga atttagaaga aggggtgcct ccatttcagc tctgccttca ctacagagac 2040

tttattcccg gtgtggccat tgctgccaac atcatccatg aaggtttcca taaaagccga 2100

aaggtgattg ttgtggtgtc ccagcacttc atccagagcc gctggtgtat ctttgaatat 2160

gagattgctc agacctggca gtttctgagc agtcgtgctg gtatcatctt cattgtcctg 2220

cagaaggtgg agaagaccct gctcaggcgg caggtggagc tgtaccgcct tctcagcagg 2280

aacacttacc tggagtggga ggacagtgtc ctggggcggc acatcttctg gagacgactc 2340

agaaaagccc tgctggatgg taaatcatgg aatccagaag gaacagtggg tacaggatgc 2400

aattggcagg aagcaacatc tatctga 2427

<210> 2
<211> 2427
<212> DNA
<213> Pan troglodytes

<220>
<221> CDS
<222> (1)..(2427)

<400> 2
gtg gtt cct aat att act tat caa tgc atg gag ctg aat ttc tac aaa 48
Val Val Pro Asn Ile Thr Tyr Gln Cys Met Glu Leu Asn Phe Tyr Lys

```
1               5                   10                  15

atc ccc gac aac ctc ccc ttc tca acc aag aac ctg gac ctg agc ttt      96
Ile Pro Asp Asn Leu Pro Phe Ser Thr Lys Asn Leu Asp Leu Ser Phe
            20                  25                  30

aat ccc ctg agg cat tta ggc agc tat agc ttc ttc agt ttc cca gaa     144
Asn Pro Leu Arg His Leu Gly Ser Tyr Ser Phe Phe Ser Phe Pro Glu
        35                  40                  45

ctg cag gtg ctg gat tta tcc agg tgt gaa atc cag aca att gaa gat     192
Leu Gln Val Leu Asp Leu Ser Arg Cys Glu Ile Gln Thr Ile Glu Asp
    50                  55                  60

ggg gca tat cag agc cta agc cac ctc tcc acc tta ata ttg aca gga     240
Gly Ala Tyr Gln Ser Leu Ser His Leu Ser Thr Leu Ile Leu Thr Gly
65                  70                  75                  80

aac ccc atc cag agt tta gcc ctg gga gcc ttt tct gga cta tca agt     288
Asn Pro Ile Gln Ser Leu Ala Leu Gly Ala Phe Ser Gly Leu Ser Ser
                85                  90                  95

tta cag aag ctg gtg gct gtg gag aca aat cta gca tct cta gag aac     336
Leu Gln Lys Leu Val Ala Val Glu Thr Asn Leu Ala Ser Leu Glu Asn
            100                 105                 110

ttc ccc att gga cat ctc aaa act ttg aaa gaa ctt aat gtg gct cac     384
Phe Pro Ile Gly His Leu Lys Thr Leu Lys Glu Leu Asn Val Ala His
        115                 120                 125

aat ctt atc caa tct ttc aaa tta cct gag tat ttt tct aat ctg acc     432
Asn Leu Ile Gln Ser Phe Lys Leu Pro Glu Tyr Phe Ser Asn Leu Thr
    130                 135                 140

aat cta gag cac ttg gac ctt tcc agc aac aag att caa agt att tat     480
Asn Leu Glu His Leu Asp Leu Ser Ser Asn Lys Ile Gln Ser Ile Tyr
145                 150                 155                 160

tgc aca gac ttg cgg gtt cta cat caa atg ccc cta ctc aat ctc tct     528
Cys Thr Asp Leu Arg Val Leu His Gln Met Pro Leu Leu Asn Leu Ser
                165                 170                 175

tta gac ctg tcc ctg aac cct atg aac ttt atc caa cca ggt gca ttt     576
Leu Asp Leu Ser Leu Asn Pro Met Asn Phe Ile Gln Pro Gly Ala Phe
            180                 185                 190

aaa gaa att agg ctt cat aag ctg act ttg aga aat aat ttt gat agt     624
Lys Glu Ile Arg Leu His Lys Leu Thr Leu Arg Asn Asn Phe Asp Ser
        195                 200                 205

tta aat gta atg aaa act tgt att caa ggt ctg gct ggt tta gaa gtc     672
Leu Asn Val Met Lys Thr Cys Ile Gln Gly Leu Ala Gly Leu Glu Val
    210                 215                 220

cat cgt ttg gtt ctg gga gaa ttt aga aat gaa gga aac ttg gaa aag     720
His Arg Leu Val Leu Gly Glu Phe Arg Asn Glu Gly Asn Leu Glu Lys
225                 230                 235                 240

ttt gac aaa tct gct cta gag ggc ctg tgc aat ttg acc att gaa gaa     768
Phe Asp Lys Ser Ala Leu Glu Gly Leu Cys Asn Leu Thr Ile Glu Glu
```

```
                245                 250                 255

ttc cga tta gca tac tta gac tac tac ctc gat gat att att gac tta      816
Phe Arg Leu Ala Tyr Leu Asp Tyr Tyr Leu Asp Asp Ile Ile Asp Leu
            260                 265                 270

ttt aat tgt ttg aca aat gtt tct tca ttt tcc ctg gtg agt gtg act      864
Phe Asn Cys Leu Thr Asn Val Ser Ser Phe Ser Leu Val Ser Val Thr
        275                 280                 285

att aaa agc gta aaa gac ttt tct tat aat ttc gga tgg caa cat tta      912
Ile Lys Ser Val Lys Asp Phe Ser Tyr Asn Phe Gly Trp Gln His Leu
    290                 295                 300

gaa tta gtt aac tgt aaa ttt gga cag ttt ccc aca ttg aaa ctc aaa      960
Glu Leu Val Asn Cys Lys Phe Gly Gln Phe Pro Thr Leu Lys Leu Lys
305                 310                 315                 320

tct ctc aaa agg ctt act ttc act tcc aac aaa ggt ggg aat gct ttt     1008
Ser Leu Lys Arg Leu Thr Phe Thr Ser Asn Lys Gly Gly Asn Ala Phe
                325                 330                 335

tca gaa gtt gat cta cca agc ctt gag ttt cta gat ctc agt aga aat     1056
Ser Glu Val Asp Leu Pro Ser Leu Glu Phe Leu Asp Leu Ser Arg Asn
            340                 345                 350

ggc ttg agt ttc aaa ggt tgc tgt tct caa agt gat ttt ggg aca acc     1104
Gly Leu Ser Phe Lys Gly Cys Cys Ser Gln Ser Asp Phe Gly Thr Thr
        355                 360                 365

agc cta aag tat tta gat ctg agc ttc aat ggt gtt att acc atg agt     1152
Ser Leu Lys Tyr Leu Asp Leu Ser Phe Asn Gly Val Ile Thr Met Ser
    370                 375                 380

tca aac ttc ttg ggc tta gaa caa cta gaa cat ctg gat ttc cag cat     1200
Ser Asn Phe Leu Gly Leu Glu Gln Leu Glu His Leu Asp Phe Gln His
385                 390                 395                 400

tcc aat ttg aaa caa atg agt gag ttt tca gta ttc cta tca ctc aga     1248
Ser Asn Leu Lys Gln Met Ser Glu Phe Ser Val Phe Leu Ser Leu Arg
                405                 410                 415

aac ctc att tac ctt gac att tct cat act cac acc aga gtt gct ttc     1296
Asn Leu Ile Tyr Leu Asp Ile Ser His Thr His Thr Arg Val Ala Phe
            420                 425                 430

aat ggc atc ttc aat ggc ttg tcc agt ctc gaa gtc ttg aaa atg gct     1344
Asn Gly Ile Phe Asn Gly Leu Ser Ser Leu Glu Val Leu Lys Met Ala
        435                 440                 445

ggc aat tct ttc cag gaa aac ttc ctt cca gat atc ttc aca gag ctg     1392
Gly Asn Ser Phe Gln Glu Asn Phe Leu Pro Asp Ile Phe Thr Glu Leu
    450                 455                 460

aga aac ttg acc ttc ctg gac ctc tct cag tgt caa ctg gag cag ttg     1440
Arg Asn Leu Thr Phe Leu Asp Leu Ser Gln Cys Gln Leu Glu Gln Leu
465                 470                 475                 480

tct cca aca gca ttt aac tca ctc tcc agt ctt cag gta cta aat atg     1488
Ser Pro Thr Ala Phe Asn Ser Leu Ser Ser Leu Gln Val Leu Asn Met
```

```
                485                 490                 495

agc cac aac aac ttc ttt tca ttg gat acg ttt cct tat aag tgt ctg     1536
Ser His Asn Asn Phe Phe Ser Leu Asp Thr Phe Pro Tyr Lys Cys Leu
            500                 505                 510

aac tcc ctc cag gtt ctt gat tac agt ctc aat cac ata atg act tcc     1584
Asn Ser Leu Gln Val Leu Asp Tyr Ser Leu Asn His Ile Met Thr Ser
        515                 520                 525

aaa aaa cag gaa cta cag cat ttt cca agt agt cta gct ttc tta aat     1632
Lys Lys Gln Glu Leu Gln His Phe Pro Ser Ser Leu Ala Phe Leu Asn
    530                 535                 540

ctt act cag aat gac ttt gct tgt act tgt gaa cac caa agt ttc ctg     1680
Leu Thr Gln Asn Asp Phe Ala Cys Thr Cys Glu His Gln Ser Phe Leu
545                 550                 555                 560

caa tgg atc aag gac cag agg cag ctc ttg gtg gaa gtt gaa cga atg     1728
Gln Trp Ile Lys Asp Gln Arg Gln Leu Leu Val Glu Val Glu Arg Met
                565                 570                 575

gaa tgt gca aca cct tca gat aag cag ggc atg cct gtg ctg agt ttg     1776
Glu Cys Ala Thr Pro Ser Asp Lys Gln Gly Met Pro Val Leu Ser Leu
            580                 585                 590

aat atc acc tgt cag atg aat aag acc atc att ggt gtg tcg gtc ctc     1824
Asn Ile Thr Cys Gln Met Asn Lys Thr Ile Ile Gly Val Ser Val Leu
        595                 600                 605

agt gtg ctt gta gta tct gtt gta gca gtt ctg gtc tat aag ttc tat     1872
Ser Val Leu Val Val Ser Val Val Ala Val Leu Val Tyr Lys Phe Tyr
    610                 615                 620

ttt cac ctg atg ctt ctt gct ggc tgc ata aag tat ggt aga ggt gaa     1920
Phe His Leu Met Leu Leu Ala Gly Cys Ile Lys Tyr Gly Arg Gly Glu
625                 630                 635                 640

aac atc tat gat gcc ttt gtt atc tac tca agc cag gat gag gac tgg     1968
Asn Ile Tyr Asp Ala Phe Val Ile Tyr Ser Ser Gln Asp Glu Asp Trp
                645                 650                 655

gta agg aat gag cta gta aag aat tta gaa gaa ggg gtg cct cca ttt     2016
Val Arg Asn Glu Leu Val Lys Asn Leu Glu Glu Gly Val Pro Pro Phe
            660                 665                 670

cag ctc tgc ctt cac tac aga gac ttt att ccc ggt gtg gcc att gct     2064
Gln Leu Cys Leu His Tyr Arg Asp Phe Ile Pro Gly Val Ala Ile Ala
        675                 680                 685

gcc aac atc atc cat gaa ggt ttc cat aaa agc cga aag gtg att gtt     2112
Ala Asn Ile Ile His Glu Gly Phe His Lys Ser Arg Lys Val Ile Val
    690                 695                 700

gtg gtg tcc cag cac ttc atc cag agc cgc tgg tgt atc ttt gaa tat     2160
Val Val Ser Gln His Phe Ile Gln Ser Arg Trp Cys Ile Phe Glu Tyr
705                 710                 715                 720

gag att gct cag acc tgg cag ttt ctg agc agt cgt gct ggt atc atc     2208
Glu Ile Ala Gln Thr Trp Gln Phe Leu Ser Ser Arg Ala Gly Ile Ile
```

```
                725                 730                 735

ttc att gtc ctg cag aag gtg gag aag acc ctg ctc agg cgg cag gtg     2256
Phe Ile Val Leu Gln Lys Val Glu Lys Thr Leu Leu Arg Arg Gln Val
            740                 745                 750

gag ctg tac cgc ctt ctc agc agg aac act tac ctg gag tgg gag gac     2304
Glu Leu Tyr Arg Leu Leu Ser Arg Asn Thr Tyr Leu Glu Trp Glu Asp
        755                 760                 765

agt gtc ctg ggg cgg cac atc ttc tgg aga cga ctc aga aaa gcc ctg     2352
Ser Val Leu Gly Arg His Ile Phe Trp Arg Arg Leu Arg Lys Ala Leu
    770                 775                 780

ctg gat ggt aaa tca tgg aat cca gaa gga aca gtg ggt aca gga tgc     2400
Leu Asp Gly Lys Ser Trp Asn Pro Glu Gly Thr Val Gly Thr Gly Cys
785                 790                 795                 800

aat tgg cag gaa gca aca tct atc tga                                2427
Asn Trp Gln Glu Ala Thr Ser Ile
                805
```

<210> 3
<211> 808
<212> PRT
<213> Pan troglodytes

<400> 3

```
Val Val Pro Asn Ile Thr Tyr Gln Cys Met Glu Leu Asn Phe Tyr Lys
1               5                   10                  15

Ile Pro Asp Asn Leu Pro Phe Ser Thr Lys Asn Leu Asp Leu Ser Phe
            20                  25                  30

Asn Pro Leu Arg His Leu Gly Ser Tyr Ser Phe Phe Ser Phe Pro Glu
        35                  40                  45

Leu Gln Val Leu Asp Leu Ser Arg Cys Glu Ile Gln Thr Ile Glu Asp
    50                  55                  60

Gly Ala Tyr Gln Ser Leu Ser His Leu Ser Thr Leu Ile Leu Thr Gly
65                  70                  75                  80

Asn Pro Ile Gln Ser Leu Ala Leu Gly Ala Phe Ser Gly Leu Ser Ser
                85                  90                  95

Leu Gln Lys Leu Val Ala Val Glu Thr Asn Leu Ala Ser Leu Glu Asn
            100                 105                 110

Phe Pro Ile Gly His Leu Lys Thr Leu Lys Glu Leu Asn Val Ala His
```

115 120 125

Asn Leu Ile Gln Ser Phe Lys Leu Pro Glu Tyr Phe Ser Asn Leu Thr
130 135 140

Asn Leu Glu His Leu Asp Leu Ser Ser Asn Lys Ile Gln Ser Ile Tyr
145 150 155 160

Cys Thr Asp Leu Arg Val Leu His Gln Met Pro Leu Leu Asn Leu Ser
165 170 175

Leu Asp Leu Ser Leu Asn Pro Met Asn Phe Ile Gln Pro Gly Ala Phe
180 185 190

Lys Glu Ile Arg Leu His Lys Leu Thr Leu Arg Asn Asn Phe Asp Ser
195 200 205

Leu Asn Val Met Lys Thr Cys Ile Gln Gly Leu Ala Gly Leu Glu Val
210 215 220

His Arg Leu Val Leu Gly Glu Phe Arg Asn Glu Gly Asn Leu Glu Lys
225 230 235 240

Phe Asp Lys Ser Ala Leu Glu Gly Leu Cys Asn Leu Thr Ile Glu Glu
245 250 255

Phe Arg Leu Ala Tyr Leu Asp Tyr Tyr Leu Asp Asp Ile Ile Asp Leu
260 265 270

Phe Asn Cys Leu Thr Asn Val Ser Ser Phe Ser Leu Val Ser Val Thr
275 280 285

Ile Lys Ser Val Lys Asp Phe Ser Tyr Asn Phe Gly Trp Gln His Leu
290 295 300

Glu Leu Val Asn Cys Lys Phe Gly Gln Phe Pro Thr Leu Lys Leu Lys
305 310 315 320

Ser Leu Lys Arg Leu Thr Phe Thr Ser Asn Lys Gly Gly Asn Ala Phe
325 330 335

Ser Glu Val Asp Leu Pro Ser Leu Glu Phe Leu Asp Leu Ser Arg Asn
340 345 350

Gly Leu Ser Phe Lys Gly Cys Cys Ser Gln Ser Asp Phe Gly Thr Thr

355 360 365

Ser Leu Lys Tyr Leu Asp Leu Ser Phe Asn Gly Val Ile Thr Met Ser
370 375 380

Ser Asn Phe Leu Gly Leu Glu Gln Leu Glu His Leu Asp Phe Gln His
385 390 395 400

Ser Asn Leu Lys Gln Met Ser Glu Phe Ser Val Phe Leu Ser Leu Arg
405 410 415

Asn Leu Ile Tyr Leu Asp Ile Ser His Thr His Thr Arg Val Ala Phe
420 425 430

Asn Gly Ile Phe Asn Gly Leu Ser Ser Leu Glu Val Leu Lys Met Ala
435 440 445

Gly Asn Ser Phe Gln Glu Asn Phe Leu Pro Asp Ile Phe Thr Glu Leu
450 455 460

Arg Asn Leu Thr Phe Leu Asp Leu Ser Gln Cys Gln Leu Glu Gln Leu
465 470 475 480

Ser Pro Thr Ala Phe Asn Ser Leu Ser Ser Leu Gln Val Leu Asn Met
485 490 495

Ser His Asn Asn Phe Phe Ser Leu Asp Thr Phe Pro Tyr Lys Cys Leu
500 505 510

Asn Ser Leu Gln Val Leu Asp Tyr Ser Leu Asn His Ile Met Thr Ser
515 520 525

Lys Lys Gln Glu Leu Gln His Phe Pro Ser Ser Leu Ala Phe Leu Asn
530 535 540

Leu Thr Gln Asn Asp Phe Ala Cys Thr Cys Glu His Gln Ser Phe Leu
545 550 555 560

Gln Trp Ile Lys Asp Gln Arg Gln Leu Leu Val Glu Val Glu Arg Met
565 570 575

Glu Cys Ala Thr Pro Ser Asp Lys Gln Gly Met Pro Val Leu Ser Leu
580 585 590

Asn Ile Thr Cys Gln Met Asn Lys Thr Ile Ile Gly Val Ser Val Leu
595 600 605

Ser Val Leu Val Val Ser Val Val Ala Val Leu Val Tyr Lys Phe Tyr
610 615 620

Phe His Leu Met Leu Leu Ala Gly Cys Ile Lys Tyr Gly Arg Gly Glu
625 630 635 640

Asn Ile Tyr Asp Ala Phe Val Ile Tyr Ser Ser Gln Asp Glu Asp Trp
645 650 655

Val Arg Asn Glu Leu Val Lys Asn Leu Glu Glu Gly Val Pro Pro Phe
660 665 670

Gln Leu Cys Leu His Tyr Arg Asp Phe Ile Pro Gly Val Ala Ile Ala
675 680 685

Ala Asn Ile Ile His Glu Gly Phe His Lys Ser Arg Lys Val Ile Val
690 695 700

Val Val Ser Gln His Phe Ile Gln Ser Arg Trp Cys Ile Phe Glu Tyr
705 710 715 720

Glu Ile Ala Gln Thr Trp Gln Phe Leu Ser Ser Arg Ala Gly Ile Ile
725 730 735

Phe Ile Val Leu Gln Lys Val Glu Lys Thr Leu Leu Arg Arg Gln Val
740 745 750

Glu Leu Tyr Arg Leu Leu Ser Arg Asn Thr Tyr Leu Glu Trp Glu Asp
755 760 765

Ser Val Leu Gly Arg His Ile Phe Trp Arg Arg Leu Arg Lys Ala Leu
770 775 780

Leu Asp Gly Lys Ser Trp Asn Pro Glu Gly Thr Val Gly Thr Gly Cys
785 790 795 800

Asn Trp Gln Glu Ala Thr Ser Ile
805

<210> 4
<211> 2427
<212> DNA
<213> Gorilla gorilla

<400> 4

gtggttccta atattactta tcaatgcatg gagctgaatt tctacaaaat ccccgacaac 60

ctccccttct caaccaagaa cctggacctg agctttaatc ccctgaggca tttaggcagc 120

tatagcttct tcagtttccc agaactgcag gtgctggatt tatccaggtg tgaaatccag 180

acaattgaag atggggcata tcagagccta agccacctct ccaccttaat attgacagga 240

aaccccatcc agagtttagc cctgggagcc ttttctggac tatcaagttt acagaagctg 300

gtggctgtgg agacaaatct agcatctcta gagaacttcc ccattggaca tctcaaaact 360

ttgaaagaac ttaatgtggc tcacaatctt attcaatctt tcaaattacc tgagtatttt 420

tctaatctga ccaatctaga gtacttggac ctttccagca acaagattca aagtatttat 480

tgcacagact tgcgggttct acatcaaatg cccctactca atctctcttt agacctgtcc 540

ctgaacccta tgacctttat ccaaccaggt gcatttaaag aaattaggct tcataagctg 600

actttgagaa ataattttga tagtttaaat gtaatgaaaa cttgtattca aggtctggct 660

ggtttagaag tccgtcgttt ggttctggga gaatttagaa atgaaggaaa cttggaaaag 720

tttgacaaat ctgctctaga gggcctgtgc aatttgacca ttgaagaatt ccgattagca 780

tacttagact actacctcga tgatattatt gacttattta attgtttgac aaatgtttct 840

tcattttccc tggtgagtgt gactattgaa agggtaaaag acttttctta taatttcgga 900

tggcaacatt tagaattagt taactgtaaa tttggacagt ttcccacatt gaaactcaaa 960

tctctcaaaa ggcttacttt cacttccaac aaaggtggga atgctttttc ggaagttgat 1020

ctaccaagcc ttgagtttct agatctcagt agaaatggct tgagtttcaa aggttgctgt 1080

tctcaaagtg attttgggac aaccagccta aagtatttag atctgagctt caatggtgtt 1140

attaccatga gttcaaactt cttgggctta gaacaactag aacatctgga tttccagcat 1200

tccaatttga aacaaatgag tgagttttca gtattcctat cactcagaaa cctcatttac 1260

cttgacattt ctcatactca caccagagtt gctttcaatg gcatcttcaa tggcttgtcc 1320

agtctcgaag tcttgaaaat ggctggcaat tctttccagg aaaacttcct tccagatatc 1380

ttcacagagc tgagaaactt gaccttcctg gacctctctc agtgtcaact ggagcagttg 1440

tctccaacag catttaactc actctccagt cttcaggtac taaatatgag ccacaacaac 1500

ttcttttcat tggatacgtt tccttataag tgtctgaact ccctccgggt tcttgattac 1560

agtctcaatc acataatgac ttccaaaaaa caggaactac agcattttcc aagcagtcta 1620

gctttcttaa atcttactca gaatgacttt gcttgtactt gtgaacacca gagtttcctg 1680

caatggatca aggaccagag gcagctcttg gtggaagttg aacgaatgga atgtgcaaca 1740

ccttcagata agcagggcat gcctgtgctg agtttgaata tcacctgtca gatgaataag 1800

accatcattg gtgtgtcggt cctcagtgtg cttgtagtat ctgttgtagc agttctggtc 1860

tataagttct attttcacct gatgcttctt gctggctgca taaagtatgg tagaggtgaa 1920

aacgtctatg atgcctttgt tatctactca agccaggatg aggactgggt aaggaatgag 1980

ctagtaaaga atttagaaga aggggtgcct ccatttcagc tctgccttca ctacagagac 2040

tttattcccg gtgtggccat tgctgccaac atcatccatg aaggtttcca taaaagtcga 2100

aaggtgattg ttgtggtgtc ccagcacttc atccagagcc gctggtgtat ctttgaatat 2160

gagattgctc agacctggca gtttctgagc agtcgtgctg gtatcatctt cattgtcctg 2220

cagaaggtgg agaagaccct gctcaggcag caggtggagc tgtaccgcct tctcagcagg 2280

aacacttacc tggagtggga ggacagtgtc ctggggcggc acatcttctg gagacgactc 2340

agaaaagccc tgctggatgg taaatcatgg aatccagaag gaacagtggg tacaggatgc 2400

aattggcagg aagcaacatc tatctga 2427

<210> 5
<211> 2427
<212> DNA
<213> Gorilla gorilla

<220>
<221> CDS
<222> (1)..(2427)

<400> 5

```
gtg gtt cct aat att act tat caa tgc atg gag ctg aat ttc tac aaa      48
Val Val Pro Asn Ile Thr Tyr Gln Cys Met Glu Leu Asn Phe Tyr Lys
1               5                  10                  15

atc ccc gac aac ctc ccc ttc tca acc aag aac ctg gac ctg agc ttt      96
Ile Pro Asp Asn Leu Pro Phe Ser Thr Lys Asn Leu Asp Leu Ser Phe
            20                  25                  30

aat ccc ctg agg cat tta ggc agc tat agc ttc ttc agt ttc cca gaa     144
Asn Pro Leu Arg His Leu Gly Ser Tyr Ser Phe Phe Ser Phe Pro Glu
        35                  40                  45

ctg cag gtg ctg gat tta tcc agg tgt gaa atc cag aca att gaa gat     192
Leu Gln Val Leu Asp Leu Ser Arg Cys Glu Ile Gln Thr Ile Glu Asp
    50                  55                  60

ggg gca tat cag agc cta agc cac ctc tcc acc tta ata ttg aca gga     240
Gly Ala Tyr Gln Ser Leu Ser His Leu Ser Thr Leu Ile Leu Thr Gly
65                  70                  75                  80

aac ccc atc cag agt tta gcc ctg gga gcc ttt tct gga cta tca agt     288
Asn Pro Ile Gln Ser Leu Ala Leu Gly Ala Phe Ser Gly Leu Ser Ser
                85                  90                  95
```

```
tta cag aag ctg gtg gct gtg gag aca aat cta gca tct cta gag aac      336
Leu Gln Lys Leu Val Ala Val Glu Thr Asn Leu Ala Ser Leu Glu Asn
            100                 105                 110

ttc ccc att gga cat ctc aaa act ttg aaa gaa ctt aat gtg gct cac      384
Phe Pro Ile Gly His Leu Lys Thr Leu Lys Glu Leu Asn Val Ala His
        115                 120                 125

aat ctt att caa tct ttc aaa tta cct gag tat ttt tct aat ctg acc      432
Asn Leu Ile Gln Ser Phe Lys Leu Pro Glu Tyr Phe Ser Asn Leu Thr
    130                 135                 140

aat cta gag tac ttg gac ctt tcc agc aac aag att caa agt att tat      480
Asn Leu Glu Tyr Leu Asp Leu Ser Ser Asn Lys Ile Gln Ser Ile Tyr
145                 150                 155                 160

tgc aca gac ttg cgg gtt cta cat caa atg ccc cta ctc aat ctc tct      528
Cys Thr Asp Leu Arg Val Leu His Gln Met Pro Leu Leu Asn Leu Ser
                165                 170                 175

tta gac ctg tcc ctg aac cct atg acc ttt atc caa cca ggt gca ttt      576
Leu Asp Leu Ser Leu Asn Pro Met Thr Phe Ile Gln Pro Gly Ala Phe
            180                 185                 190

aaa gaa att agg ctt cat aag ctg act ttg aga aat aat ttt gat agt      624
Lys Glu Ile Arg Leu His Lys Leu Thr Leu Arg Asn Asn Phe Asp Ser
        195                 200                 205

tta aat gta atg aaa act tgt att caa ggt ctg gct ggt tta gaa gtc      672
Leu Asn Val Met Lys Thr Cys Ile Gln Gly Leu Ala Gly Leu Glu Val
    210                 215                 220

cgt cgt ttg gtt ctg gga gaa ttt aga aat gaa gga aac ttg gaa aag      720
Arg Arg Leu Val Leu Gly Glu Phe Arg Asn Glu Gly Asn Leu Glu Lys
225                 230                 235                 240

ttt gac aaa tct gct cta gag ggc ctg tgc aat ttg acc att gaa gaa      768
Phe Asp Lys Ser Ala Leu Glu Gly Leu Cys Asn Leu Thr Ile Glu Glu
                245                 250                 255

ttc cga tta gca tac tta gac tac tac ctc gat gat att att gac tta      816
Phe Arg Leu Ala Tyr Leu Asp Tyr Tyr Leu Asp Asp Ile Ile Asp Leu
            260                 265                 270

ttt aat tgt ttg aca aat gtt tct tca ttt tcc ctg gtg agt gtg act      864
Phe Asn Cys Leu Thr Asn Val Ser Ser Phe Ser Leu Val Ser Val Thr
        275                 280                 285

att gaa agg gta aaa gac ttt tct tat aat ttc gga tgg caa cat tta      912
Ile Glu Arg Val Lys Asp Phe Ser Tyr Asn Phe Gly Trp Gln His Leu
    290                 295                 300

gaa tta gtt aac tgt aaa ttt gga cag ttt ccc aca ttg aaa ctc aaa      960
Glu Leu Val Asn Cys Lys Phe Gly Gln Phe Pro Thr Leu Lys Leu Lys
305                 310                 315                 320

tct ctc aaa agg ctt act ttc act tcc aac aaa ggt ggg aat gct ttt     1008
Ser Leu Lys Arg Leu Thr Phe Thr Ser Asn Lys Gly Gly Asn Ala Phe
                325                 330                 335
```

```
tcg gaa gtt gat cta cca agc ctt gag ttt cta gat ctc agt aga aat      1056
Ser Glu Val Asp Leu Pro Ser Leu Glu Phe Leu Asp Leu Ser Arg Asn
            340                 345                 350

ggc ttg agt ttc aaa ggt tgc tgt tct caa agt gat ttt ggg aca acc      1104
Gly Leu Ser Phe Lys Gly Cys Cys Ser Gln Ser Asp Phe Gly Thr Thr
        355                 360                 365

agc cta aag tat tta gat ctg agc ttc aat ggt gtt att acc atg agt      1152
Ser Leu Lys Tyr Leu Asp Leu Ser Phe Asn Gly Val Ile Thr Met Ser
    370                 375                 380

tca aac ttc ttg ggc tta gaa caa cta gaa cat ctg gat ttc cag cat      1200
Ser Asn Phe Leu Gly Leu Glu Gln Leu Glu His Leu Asp Phe Gln His
385                 390                 395                 400

tcc aat ttg aaa caa atg agt gag ttt tca gta ttc cta tca ctc aga      1248
Ser Asn Leu Lys Gln Met Ser Glu Phe Ser Val Phe Leu Ser Leu Arg
                405                 410                 415

aac ctc att tac ctt gac att tct cat act cac acc aga gtt gct ttc      1296
Asn Leu Ile Tyr Leu Asp Ile Ser His Thr His Thr Arg Val Ala Phe
            420                 425                 430

aat ggc atc ttc aat ggc ttg tcc agt ctc gaa gtc ttg aaa atg gct      1344
Asn Gly Ile Phe Asn Gly Leu Ser Ser Leu Glu Val Leu Lys Met Ala
        435                 440                 445

ggc aat tct ttc cag gaa aac ttc ctt cca gat atc ttc aca gag ctg      1392
Gly Asn Ser Phe Gln Glu Asn Phe Leu Pro Asp Ile Phe Thr Glu Leu
    450                 455                 460

aga aac ttg acc ttc ctg gac ctc tct cag tgt caa ctg gag cag ttg      1440
Arg Asn Leu Thr Phe Leu Asp Leu Ser Gln Cys Gln Leu Glu Gln Leu
465                 470                 475                 480

tct cca aca gca ttt aac tca ctc tcc agt ctt cag gta cta aat atg      1488
Ser Pro Thr Ala Phe Asn Ser Leu Ser Ser Leu Gln Val Leu Asn Met
                485                 490                 495

agc cac aac aac ttc ttt tca ttg gat acg ttt cct tat aag tgt ctg      1536
Ser His Asn Asn Phe Phe Ser Leu Asp Thr Phe Pro Tyr Lys Cys Leu
            500                 505                 510

aac tcc ctc cgg gtt ctt gat tac agt ctc aat cac ata atg act tcc      1584
Asn Ser Leu Arg Val Leu Asp Tyr Ser Leu Asn His Ile Met Thr Ser
        515                 520                 525

aaa aaa cag gaa cta cag cat ttt cca agc agt cta gct ttc tta aat      1632
Lys Lys Gln Glu Leu Gln His Phe Pro Ser Ser Leu Ala Phe Leu Asn
    530                 535                 540

ctt act cag aat gac ttt gct tgt act tgt gaa cac cag agt ttc ctg      1680
Leu Thr Gln Asn Asp Phe Ala Cys Thr Cys Glu His Gln Ser Phe Leu
545                 550                 555                 560

caa tgg atc aag gac cag agg cag ctc ttg gtg gaa gtt gaa cga atg      1728
Gln Trp Ile Lys Asp Gln Arg Gln Leu Leu Val Glu Val Glu Arg Met
                565                 570                 575
```

```
gaa tgt gca aca cct tca gat aag cag ggc atg cct gtg ctg agt ttg     1776
Glu Cys Ala Thr Pro Ser Asp Lys Gln Gly Met Pro Val Leu Ser Leu
            580                 585                 590

aat atc acc tgt cag atg aat aag acc atc att ggt gtg tcg gtc ctc     1824
Asn Ile Thr Cys Gln Met Asn Lys Thr Ile Ile Gly Val Ser Val Leu
        595                 600                 605

agt gtg ctt gta gta tct gtt gta gca gtt ctg gtc tat aag ttc tat     1872
Ser Val Leu Val Val Ser Val Val Ala Val Leu Val Tyr Lys Phe Tyr
    610                 615                 620

ttt cac ctg atg ctt ctt gct ggc tgc ata aag tat ggt aga ggt gaa     1920
Phe His Leu Met Leu Leu Ala Gly Cys Ile Lys Tyr Gly Arg Gly Glu
625                 630                 635                 640

aac gtc tat gat gcc ttt gtt atc tac tca agc cag gat gag gac tgg     1968
Asn Val Tyr Asp Ala Phe Val Ile Tyr Ser Ser Gln Asp Glu Asp Trp
                645                 650                 655

gta agg aat gag cta gta aag aat tta gaa gaa ggg gtg cct cca ttt     2016
Val Arg Asn Glu Leu Val Lys Asn Leu Glu Glu Gly Val Pro Pro Phe
            660                 665                 670

cag ctc tgc ctt cac tac aga gac ttt att ccc ggt gtg gcc att gct     2064
Gln Leu Cys Leu His Tyr Arg Asp Phe Ile Pro Gly Val Ala Ile Ala
        675                 680                 685

gcc aac atc atc cat gaa ggt ttc cat aaa agt cga aag gtg att gtt     2112
Ala Asn Ile Ile His Glu Gly Phe His Lys Ser Arg Lys Val Ile Val
    690                 695                 700

gtg gtg tcc cag cac ttc atc cag agc cgc tgg tgt atc ttt gaa tat     2160
Val Val Ser Gln His Phe Ile Gln Ser Arg Trp Cys Ile Phe Glu Tyr
705                 710                 715                 720

gag att gct cag acc tgg cag ttt ctg agc agt cgt gct ggt atc atc     2208
Glu Ile Ala Gln Thr Trp Gln Phe Leu Ser Ser Arg Ala Gly Ile Ile
                725                 730                 735

ttc att gtc ctg cag aag gtg gag aag acc ctg ctc agg cag cag gtg     2256
Phe Ile Val Leu Gln Lys Val Glu Lys Thr Leu Leu Arg Gln Gln Val
            740                 745                 750

gag ctg tac cgc ctt ctc agc agg aac act tac ctg gag tgg gag gac     2304
Glu Leu Tyr Arg Leu Leu Ser Arg Asn Thr Tyr Leu Glu Trp Glu Asp
        755                 760                 765

agt gtc ctg ggg cgg cac atc ttc tgg aga cga ctc aga aaa gcc ctg     2352
Ser Val Leu Gly Arg His Ile Phe Trp Arg Arg Leu Arg Lys Ala Leu
    770                 775                 780

ctg gat ggt aaa tca tgg aat cca gaa gga aca gtg ggt aca gga tgc     2400
Leu Asp Gly Lys Ser Trp Asn Pro Glu Gly Thr Val Gly Thr Gly Cys
785                 790                 795                 800

aat tgg cag gaa gca aca tct atc tga                                2427
Asn Trp Gln Glu Ala Thr Ser Ile
                805
```

<210> 6
<211> 808
<212> PRT
<213> Gorilla gorilla

<400> 6

Val Val Pro Asn Ile Thr Tyr Gln Cys Met Glu Leu Asn Phe Tyr Lys
1 5 10 15

Ile Pro Asp Asn Leu Pro Phe Ser Thr Lys Asn Leu Asp Leu Ser Phe
20 25 30

Asn Pro Leu Arg His Leu Gly Ser Tyr Ser Phe Phe Ser Phe Pro Glu
35 40 45

Leu Gln Val Leu Asp Leu Ser Arg Cys Glu Ile Gln Thr Ile Glu Asp
50 55 60

Gly Ala Tyr Gln Ser Leu Ser His Leu Ser Thr Leu Ile Leu Thr Gly
65 70 75 80

Asn Pro Ile Gln Ser Leu Ala Leu Gly Ala Phe Ser Gly Leu Ser Ser
85 90 95

Leu Gln Lys Leu Val Ala Val Glu Thr Asn Leu Ala Ser Leu Glu Asn
100 105 110

Phe Pro Ile Gly His Leu Lys Thr Leu Lys Glu Leu Asn Val Ala His
115 120 125

Asn Leu Ile Gln Ser Phe Lys Leu Pro Glu Tyr Phe Ser Asn Leu Thr
130 135 140

Asn Leu Glu Tyr Leu Asp Leu Ser Ser Asn Lys Ile Gln Ser Ile Tyr
145 150 155 160

Cys Thr Asp Leu Arg Val Leu His Gln Met Pro Leu Leu Asn Leu Ser
165 170 175

Leu Asp Leu Ser Leu Asn Pro Met Thr Phe Ile Gln Pro Gly Ala Phe
180 185 190

Lys Glu Ile Arg Leu His Lys Leu Thr Leu Arg Asn Asn Phe Asp Ser
195 200 205

```
Leu Asn Val Met Lys Thr Cys Ile Gln Gly Leu Ala Gly Leu Glu Val
    210                 215                 220

Arg Arg Leu Val Leu Gly Glu Phe Arg Asn Glu Gly Asn Leu Glu Lys
225                 230                 235                 240

Phe Asp Lys Ser Ala Leu Glu Gly Leu Cys Asn Leu Thr Ile Glu Glu
                245                 250                 255

Phe Arg Leu Ala Tyr Leu Asp Tyr Tyr Leu Asp Asp Ile Ile Asp Leu
            260                 265                 270

Phe Asn Cys Leu Thr Asn Val Ser Ser Phe Ser Leu Val Ser Val Thr
        275                 280                 285

Ile Glu Arg Val Lys Asp Phe Ser Tyr Asn Phe Gly Trp Gln His Leu
    290                 295                 300

Glu Leu Val Asn Cys Lys Phe Gly Gln Phe Pro Thr Leu Lys Leu Lys
305                 310                 315                 320

Ser Leu Lys Arg Leu Thr Phe Thr Ser Asn Lys Gly Gly Asn Ala Phe
                325                 330                 335

Ser Glu Val Asp Leu Pro Ser Leu Glu Phe Leu Asp Leu Ser Arg Asn
            340                 345                 350

Gly Leu Ser Phe Lys Gly Cys Cys Ser Gln Ser Asp Phe Gly Thr Thr
        355                 360                 365

Ser Leu Lys Tyr Leu Asp Leu Ser Phe Asn Gly Val Ile Thr Met Ser
    370                 375                 380

Ser Asn Phe Leu Gly Leu Glu Gln Leu Glu His Leu Asp Phe Gln His
385                 390                 395                 400

Ser Asn Leu Lys Gln Met Ser Glu Phe Ser Val Phe Leu Ser Leu Arg
                405                 410                 415

Asn Leu Ile Tyr Leu Asp Ile Ser His Thr His Thr Arg Val Ala Phe
            420                 425                 430

Asn Gly Ile Phe Asn Gly Leu Ser Ser Leu Glu Val Leu Lys Met Ala
        435                 440                 445
```

Gly Asn Ser Phe Gln Glu Asn Phe Leu Pro Asp Ile Phe Thr Glu Leu
450 455 460

Arg Asn Leu Thr Phe Leu Asp Leu Ser Gln Cys Gln Leu Glu Gln Leu
465 470 475 480

Ser Pro Thr Ala Phe Asn Ser Leu Ser Ser Leu Gln Val Leu Asn Met
485 490 495

Ser His Asn Asn Phe Phe Ser Leu Asp Thr Phe Pro Tyr Lys Cys Leu
500 505 510

Asn Ser Leu Arg Val Leu Asp Tyr Ser Leu Asn His Ile Met Thr Ser
515 520 525

Lys Lys Gln Glu Leu Gln His Phe Pro Ser Ser Leu Ala Phe Leu Asn
530 535 540

Leu Thr Gln Asn Asp Phe Ala Cys Thr Cys Glu His Gln Ser Phe Leu
545 550 555 560

Gln Trp Ile Lys Asp Gln Arg Gln Leu Leu Val Glu Val Glu Arg Met
565 570 575

Glu Cys Ala Thr Pro Ser Asp Lys Gln Gly Met Pro Val Leu Ser Leu
580 585 590

Asn Ile Thr Cys Gln Met Asn Lys Thr Ile Ile Gly Val Ser Val Leu
595 600 605

Ser Val Leu Val Val Ser Val Val Ala Val Leu Val Tyr Lys Phe Tyr
610 615 620

Phe His Leu Met Leu Leu Ala Gly Cys Ile Lys Tyr Gly Arg Gly Glu
625 630 635 640

Asn Val Tyr Asp Ala Phe Val Ile Tyr Ser Ser Gln Asp Glu Asp Trp
645 650 655

Val Arg Asn Glu Leu Val Lys Asn Leu Glu Glu Gly Val Pro Pro Phe
660 665 670

Gln Leu Cys Leu His Tyr Arg Asp Phe Ile Pro Gly Val Ala Ile Ala
675 680 685

Ala Asn Ile Ile His Glu Gly Phe His Lys Ser Arg Lys Val Ile Val
690 695 700

Val Val Ser Gln His Phe Ile Gln Ser Arg Trp Cys Ile Phe Glu Tyr
705 710 715 720

Glu Ile Ala Gln Thr Trp Gln Phe Leu Ser Ser Arg Ala Gly Ile Ile
725 730 735

Phe Ile Val Leu Gln Lys Val Glu Lys Thr Leu Leu Arg Gln Gln Val
740 745 750

Glu Leu Tyr Arg Leu Leu Ser Arg Asn Thr Tyr Leu Glu Trp Glu Asp
755 760 765

Ser Val Leu Gly Arg His Ile Phe Trp Arg Arg Leu Arg Lys Ala Leu
770 775 780

Leu Asp Gly Lys Ser Trp Asn Pro Glu Gly Thr Val Gly Thr Gly Cys
785 790 795 800

Asn Trp Gln Glu Ala Thr Ser Ile
805

<210> 7
<211> 2406
<212> DNA
<213> Hylobates lar

<400> 7
gtggttccta acattactta tcaatgcatg gagctgaatt tctacaaaat ccccgacaac 60

ctccccttct caaccaagaa cctggacctg agctttaatc ccctgaggca tttaggcagc 120

tatagcttct tcagtttccc agaactgcag gtgctggatt tatccaggtg tgaaatccag 180

acaattgaag atggggcata tcagagccta agcctcctct ccaccttaat attgacagga 240

aaccccatcc agagtttagc tctgggagcc ttttctggac tatcaagttt acagaagcta 300

gtggctgtgg agacaaatct agcatctcta gagaacttcc ccattggaca tctcaaaact 360

ttgaaagaac ttaatgtggc tcacaatctt atccaatctt tcaaattacc tgagtatttt 420

tctaatctga ccaatctaga gcacttggac ctttccagca acaagattca aagtatttat 480

tgcaaagact tgcaggttct acatcaaatg cccctactca atctctcttt agacctgtcc 540

ctgaacccta tgaactttat ccaaccaggt gcatttaaag aaattagcct tcrtaagctg 600

actttaagaa ataattttga tagtttaaat gtaatgaaaa cttgtattca aggtctggct 660

ggtttagaag tccatcgttt ggttctggga gaatttagaa atgaaggaaa cttggaagag 720

tttgacaaat ctgctctaga gggcctgtgc aatttgacca ttgaagaatt ccgattagca 780

tacttagacc actacctcga tgatattatt gacttattta attgtttggc aaatgtttct 840

tcatttttccc tggtgagtgt gactattaaa agggtagaag acttttctta taatttcgga 900

tggcaacatt tagaattagt taactgtaaa tttggacagt ttcccacatt gaacctcaaa 960

tctctcaaaa ggcttacttt cactgccaac agaggtggga atgctttttc agaagttgat 1020

ctaccaagcc ttgagtttct agatctcagt agaaatggct tgagtttcaa aggttgctgt 1080

tctcaaagtg attttgggac aaacagccta aagtatttag atctgagctt caatgatgtt 1140

attaccatga gttcaaactt cttgggctta gaacagctag aacatctgga tttgcagcat 1200

tccaatttga aacaaatgag tgaattttca gtattcctat cactcagaaa cctcatttac 1260

cttgacattt ctcatactca caccagagtt gctttcaatg gcatcttcaa tggcttgtcc 1320

aatctcgaag tcttgaaaat ggctggcaat tctttccagg aaaacttcct tccagatatc 1380

ttcacagagc tgagaaactt gaccttcctg gacctctctc agtgtcaact ggagcaattg 1440

tctccaacag catttaactc actctccagt cttcaggtac taaatatgag ccacaacaac 1500

ttcttttcat tggatacgtt tccttataag tgtctgaact ccctccaggt tcttgattac 1560

agtctcaatc acataatgac ttccaaaaaa caggaactac agcgttttcc aagtagtcta 1620

gccttcttaa atcttactca gaatgacttt gcttgtactt gtgaacacga gagtttcctg 1680

cagtggatca aggaccagag gcagctcttg gtggaagttg aacgaatgga atgtgcaaca 1740

ccttcagata agcagggcat gcctgtgctg agtttgaata tcacctgtca gatgaataag 1800

accatcattg gtgtgtcagt cctcagtgtg cttgtagtat ctgttgtagc agttctggtc 1860

tataagttct attttcacct gatgcttctt gctggctgca tgaagtatgg tagaggtgaa 1920

aacacctatg atgcctttgt tatctactcc agccaggatg aggactgggt aaggaatgag 1980

ctagtaaaga atttagaaga aggggtgcct ccctttcagc tctgccttca ctacagagac 2040

tttattccyg gtgtggccat tgctgccaac atcatccatg aaggtttcca taaaagccga 2100

aaggtgattg ttgtggtgtc ccagcacttc atccagagcc gctggtgtat ctttgagtat 2160

gagattgctc agacctggca gtttctgagc agtcatgctg ggatcatctt cattgtcctg 2220

cagaaggtgg agaagaccct gctcaggcag caggtggagc tgtaccgcct tctcagcagg 2280

aacacttacc tggagtggga ggatagtgtc ctggggcggc acattttctg gagacgactc 2340

agaaaagccc tgctggatgg taaatcatgg aatccagaag gaacagtggg tacaggatgc 2400

aattag 2406

<210> 8
<211> 2406
<212> DNA
<213> Hylobates lar

<220>
<221> CDS
<222> (1)..(2406)

<400> 8

```
gtg gtt cct aac att act tat caa tgc atg gag ctg aat ttc tac aaa      48
Val Val Pro Asn Ile Thr Tyr Gln Cys Met Glu Leu Asn Phe Tyr Lys
1               5                  10                  15

atc ccc gac aac ctc ccc ttc tca acc aag aac ctg gac ctg agc ttt      96
Ile Pro Asp Asn Leu Pro Phe Ser Thr Lys Asn Leu Asp Leu Ser Phe
            20                  25                  30

aat ccc ctg agg cat tta ggc agc tat agc ttc ttc agt ttc cca gaa     144
Asn Pro Leu Arg His Leu Gly Ser Tyr Ser Phe Phe Ser Phe Pro Glu
        35                  40                  45

ctg cag gtg ctg gat tta tcc agg tgt gaa atc cag aca att gaa gat     192
Leu Gln Val Leu Asp Leu Ser Arg Cys Glu Ile Gln Thr Ile Glu Asp
    50                  55                  60

ggg gca tat cag agc cta agc ctc ctc tcc acc tta ata ttg aca gga     240
Gly Ala Tyr Gln Ser Leu Ser Leu Leu Ser Thr Leu Ile Leu Thr Gly
65                  70                  75                  80

aac ccc atc cag agt tta gct ctg gga gcc ttt tct gga cta tca agt     288
Asn Pro Ile Gln Ser Leu Ala Leu Gly Ala Phe Ser Gly Leu Ser Ser
                85                  90                  95

tta cag aag cta gtg gct gtg gag aca aat cta gca tct cta gag aac     336
Leu Gln Lys Leu Val Ala Val Glu Thr Asn Leu Ala Ser Leu Glu Asn
            100                 105                 110

ttc ccc att gga cat ctc aaa act ttg aaa gaa ctt aat gtg gct cac     384
Phe Pro Ile Gly His Leu Lys Thr Leu Lys Glu Leu Asn Val Ala His
        115                 120                 125

aat ctt atc caa tct ttc aaa tta cct gag tat ttt tct aat ctg acc     432
Asn Leu Ile Gln Ser Phe Lys Leu Pro Glu Tyr Phe Ser Asn Leu Thr
    130                 135                 140

aat cta gag cac ttg gac ctt tcc agc aac aag att caa agt att tat     480
Asn Leu Glu His Leu Asp Leu Ser Ser Asn Lys Ile Gln Ser Ile Tyr
145                 150                 155                 160

tgc aaa gac ttg cag gtt cta cat caa atg ccc cta ctc aat ctc tct     528
Cys Lys Asp Leu Gln Val Leu His Gln Met Pro Leu Leu Asn Leu Ser
                165                 170                 175

tta gac ctg tcc ctg aac cct atg aac ttt atc caa cca ggt gca ttt     576
Leu Asp Leu Ser Leu Asn Pro Met Asn Phe Ile Gln Pro Gly Ala Phe
            180                 185                 190
```

```
aaa gaa att agc ctt crt aag ctg act tta aga aat aat ttt gat agt      624
Lys Glu Ile Ser Leu Xaa Lys Leu Thr Leu Arg Asn Asn Phe Asp Ser
        195                 200                 205

tta aat gta atg aaa act tgt att caa ggt ctg gct ggt tta gaa gtc      672
Leu Asn Val Met Lys Thr Cys Ile Gln Gly Leu Ala Gly Leu Glu Val
    210                 215                 220

cat cgt ttg gtt ctg gga gaa ttt aga aat gaa gga aac ttg gaa gag      720
His Arg Leu Val Leu Gly Glu Phe Arg Asn Glu Gly Asn Leu Glu Glu
225                 230                 235                 240

ttt gac aaa tct gct cta gag ggc ctg tgc aat ttg acc att gaa gaa      768
Phe Asp Lys Ser Ala Leu Glu Gly Leu Cys Asn Leu Thr Ile Glu Glu
                245                 250                 255

ttc cga tta gca tac tta gac cac tac ctc gat gat att att gac tta      816
Phe Arg Leu Ala Tyr Leu Asp His Tyr Leu Asp Asp Ile Ile Asp Leu
            260                 265                 270

ttt aat tgt ttg gca aat gtt tct tca ttt tcc ctg gtg agt gtg act      864
Phe Asn Cys Leu Ala Asn Val Ser Ser Phe Ser Leu Val Ser Val Thr
        275                 280                 285

att aaa agg gta gaa gac ttt tct tat aat ttc gga tgg caa cat tta      912
Ile Lys Arg Val Glu Asp Phe Ser Tyr Asn Phe Gly Trp Gln His Leu
    290                 295                 300

gaa tta gtt aac tgt aaa ttt gga cag ttt ccc aca ttg aac ctc aaa      960
Glu Leu Val Asn Cys Lys Phe Gly Gln Phe Pro Thr Leu Asn Leu Lys
305                 310                 315                 320

tct ctc aaa agg ctt act ttc act gcc aac aga ggt ggg aat gct ttt     1008
Ser Leu Lys Arg Leu Thr Phe Thr Ala Asn Arg Gly Gly Asn Ala Phe
                325                 330                 335

tca gaa gtt gat cta cca agc ctt gag ttt cta gat ctc agt aga aat     1056
Ser Glu Val Asp Leu Pro Ser Leu Glu Phe Leu Asp Leu Ser Arg Asn
            340                 345                 350

ggc ttg agt ttc aaa ggt tgc tgt tct caa agt gat ttt ggg aca aac     1104
Gly Leu Ser Phe Lys Gly Cys Cys Ser Gln Ser Asp Phe Gly Thr Asn
        355                 360                 365

agc cta aag tat tta gat ctg agc ttc aat gat gtt att acc atg agt     1152
Ser Leu Lys Tyr Leu Asp Leu Ser Phe Asn Asp Val Ile Thr Met Ser
    370                 375                 380

tca aac ttc ttg ggc tta gaa cag cta gaa cat ctg gat ttg cag cat     1200
Ser Asn Phe Leu Gly Leu Glu Gln Leu Glu His Leu Asp Leu Gln His
385                 390                 395                 400

tcc aat ttg aaa caa atg agt gaa ttt tca gta ttc cta tca ctc aga     1248
Ser Asn Leu Lys Gln Met Ser Glu Phe Ser Val Phe Leu Ser Leu Arg
                405                 410                 415

aac ctc att tac ctt gac att tct cat act cac acc aga gtt gct ttc     1296
Asn Leu Ile Tyr Leu Asp Ile Ser His Thr His Thr Arg Val Ala Phe
            420                 425                 430
```

```
aat ggc atc ttc aat ggc ttg tcc aat ctc gaa gtc ttg aaa atg gct      1344
Asn Gly Ile Phe Asn Gly Leu Ser Asn Leu Glu Val Leu Lys Met Ala
        435                 440                 445

ggc aat tct ttc cag gaa aac ttc ctt cca gat atc ttc aca gag ctg      1392
Gly Asn Ser Phe Gln Glu Asn Phe Leu Pro Asp Ile Phe Thr Glu Leu
    450                 455                 460

aga aac ttg acc ttc ctg gac ctc tct cag tgt caa ctg gag caa ttg      1440
Arg Asn Leu Thr Phe Leu Asp Leu Ser Gln Cys Gln Leu Glu Gln Leu
465                 470                 475                 480

tct cca aca gca ttt aac tca ctc tcc agt ctt cag gta cta aat atg      1488
Ser Pro Thr Ala Phe Asn Ser Leu Ser Ser Leu Gln Val Leu Asn Met
                485                 490                 495

agc cac aac aac ttc ttt tca ttg gat acg ttt cct tat aag tgt ctg      1536
Ser His Asn Asn Phe Phe Ser Leu Asp Thr Phe Pro Tyr Lys Cys Leu
            500                 505                 510

aac tcc ctc cag gtt ctt gat tac agt ctc aat cac ata atg act tcc      1584
Asn Ser Leu Gln Val Leu Asp Tyr Ser Leu Asn His Ile Met Thr Ser
        515                 520                 525

aaa aaa cag gaa cta cag cgt ttt cca agt agt cta gcc ttc tta aat      1632
Lys Lys Gln Glu Leu Gln Arg Phe Pro Ser Ser Leu Ala Phe Leu Asn
    530                 535                 540

ctt act cag aat gac ttt gct tgt act tgt gaa cac gag agt ttc ctg      1680
Leu Thr Gln Asn Asp Phe Ala Cys Thr Cys Glu His Glu Ser Phe Leu
545                 550                 555                 560

cag tgg atc aag gac cag agg cag ctc ttg gtg gaa gtt gaa cga atg      1728
Gln Trp Ile Lys Asp Gln Arg Gln Leu Leu Val Glu Val Glu Arg Met
                565                 570                 575

gaa tgt gca aca cct tca gat aag cag ggc atg cct gtg ctg agt ttg      1776
Glu Cys Ala Thr Pro Ser Asp Lys Gln Gly Met Pro Val Leu Ser Leu
            580                 585                 590

aat atc acc tgt cag atg aat aag acc atc att ggt gtg tca gtc ctc      1824
Asn Ile Thr Cys Gln Met Asn Lys Thr Ile Ile Gly Val Ser Val Leu
        595                 600                 605

agt gtg ctt gta gta tct gtt gta gca gtt ctg gtc tat aag ttc tat      1872
Ser Val Leu Val Val Ser Val Val Ala Val Leu Val Tyr Lys Phe Tyr
    610                 615                 620

ttt cac ctg atg ctt ctt gct ggc tgc atg aag tat ggt aga ggt gaa      1920
Phe His Leu Met Leu Leu Ala Gly Cys Met Lys Tyr Gly Arg Gly Glu
625                 630                 635                 640

aac acc tat gat gcc ttt gtt atc tac tcc agc cag gat gag gac tgg      1968
Asn Thr Tyr Asp Ala Phe Val Ile Tyr Ser Ser Gln Asp Glu Asp Trp
                645                 650                 655

gta agg aat gag cta gta aag aat tta gaa gaa ggg gtg cct ccc ttt      2016
Val Arg Asn Glu Leu Val Lys Asn Leu Glu Glu Gly Val Pro Pro Phe
            660                 665                 670
```

```
cag ctc tgc ctt cac tac aga gac ttt att ccy ggt gtg gcc att gct     2064
Gln Leu Cys Leu His Tyr Arg Asp Phe Ile Xaa Gly Val Ala Ile Ala
        675                 680                 685

gcc aac atc atc cat gaa ggt ttc cat aaa agc cga aag gtg att gtt     2112
Ala Asn Ile Ile His Glu Gly Phe His Lys Ser Arg Lys Val Ile Val
    690                 695                 700

gtg gtg tcc cag cac ttc atc cag agc cgc tgg tgt atc ttt gag tat     2160
Val Val Ser Gln His Phe Ile Gln Ser Arg Trp Cys Ile Phe Glu Tyr
705                 710                 715                 720

gag att gct cag acc tgg cag ttt ctg agc agt cat gct ggg atc atc     2208
Glu Ile Ala Gln Thr Trp Gln Phe Leu Ser Ser His Ala Gly Ile Ile
                725                 730                 735

ttc att gtc ctg cag aag gtg gag aag acc ctg ctc agg cag cag gtg     2256
Phe Ile Val Leu Gln Lys Val Glu Lys Thr Leu Leu Arg Gln Gln Val
            740                 745                 750

gag ctg tac cgc ctt ctc agc agg aac act tac ctg gag tgg gag gat     2304
Glu Leu Tyr Arg Leu Leu Ser Arg Asn Thr Tyr Leu Glu Trp Glu Asp
        755                 760                 765

agt gtc ctg ggg cgg cac att ttc tgg aga cga ctc aga aaa gcc ctg     2352
Ser Val Leu Gly Arg His Ile Phe Trp Arg Arg Leu Arg Lys Ala Leu
    770                 775                 780

ctg gat ggt aaa tca tgg aat cca gaa gga aca gtg ggt aca gga tgc     2400
Leu Asp Gly Lys Ser Trp Asn Pro Glu Gly Thr Val Gly Thr Gly Cys
785                 790                 795                 800

aat tag                                                            2406
Asn
```

<210> 9

<211> 801
<212> PRT
<213> Hylobates lar

<220>
<221> misc_feature
<222> (198)..(198)
<223> The 'Xaa' at location 198 stands for Arg, or His.

<220>
<221> misc_feature
<222> (683)..(683)
<223> The 'Xaa' at location 683 stands for Pro.

<400> 9

```
Val Val Pro Asn Ile Thr Tyr Gln Cys Met Glu Leu Asn Phe Tyr Lys
1               5                   10                  15
```

Ile Pro Asp Asn Leu Pro Phe Ser Thr Lys Asn Leu Asp Leu Ser Phe
20 25 30

Asn Pro Leu Arg His Leu Gly Ser Tyr Ser Phe Phe Ser Phe Pro Glu
35 40 45

Leu Gln Val Leu Asp Leu Ser Arg Cys Glu Ile Gln Thr Ile Glu Asp
50 55 60

Gly Ala Tyr Gln Ser Leu Ser Leu Leu Ser Thr Leu Ile Leu Thr Gly
65 70 75 80

Asn Pro Ile Gln Ser Leu Ala Leu Gly Ala Phe Ser Gly Leu Ser Ser
85 90 95

Leu Gln Lys Leu Val Ala Val Glu Thr Asn Leu Ala Ser Leu Glu Asn
100 105 110

Phe Pro Ile Gly His Leu Lys Thr Leu Lys Glu Leu Asn Val Ala His
115 120 125

Asn Leu Ile Gln Ser Phe Lys Leu Pro Glu Tyr Phe Ser Asn Leu Thr
130 135 140

Asn Leu Glu His Leu Asp Leu Ser Ser Asn Lys Ile Gln Ser Ile Tyr
145 150 155 160

Cys Lys Asp Leu Gln Val Leu His Gln Met Pro Leu Leu Asn Leu Ser
165 170 175

Leu Asp Leu Ser Leu Asn Pro Met Asn Phe Ile Gln Pro Gly Ala Phe
180 185 190

Lys Glu Ile Ser Leu Xaa Lys Leu Thr Leu Arg Asn Asn Phe Asp Ser
195 200 205

Leu Asn Val Met Lys Thr Cys Ile Gln Gly Leu Ala Gly Leu Glu Val
210 215 220

His Arg Leu Val Leu Gly Glu Phe Arg Asn Glu Gly Asn Leu Glu Glu
225 230 235 240

Phe Asp Lys Ser Ala Leu Glu Gly Leu Cys Asn Leu Thr Ile Glu Glu
245 250 255

Phe Arg Leu Ala Tyr Leu Asp His Tyr Leu Asp Asp Ile Ile Asp Leu
260 265 270

Phe Asn Cys Leu Ala Asn Val Ser Ser Phe Ser Leu Val Ser Val Thr
275 280 285

Ile Lys Arg Val Glu Asp Phe Ser Tyr Asn Phe Gly Trp Gln His Leu
290 295 300

Glu Leu Val Asn Cys Lys Phe Gly Gln Phe Pro Thr Leu Asn Leu Lys
305 310 315 320

Ser Leu Lys Arg Leu Thr Phe Thr Ala Asn Arg Gly Gly Asn Ala Phe
325 330 335

Ser Glu Val Asp Leu Pro Ser Leu Glu Phe Leu Asp Leu Ser Arg Asn
340 345 350

Gly Leu Ser Phe Lys Gly Cys Cys Ser Gln Ser Asp Phe Gly Thr Asn
355 360 365

Ser Leu Lys Tyr Leu Asp Leu Ser Phe Asn Asp Val Ile Thr Met Ser
370 375 380

Ser Asn Phe Leu Gly Leu Glu Gln Leu Glu His Leu Asp Leu Gln His
385 390 395 400

Ser Asn Leu Lys Gln Met Ser Glu Phe Ser Val Phe Leu Ser Leu Arg
405 410 415

Asn Leu Ile Tyr Leu Asp Ile Ser His Thr His Thr Arg Val Ala Phe
420 425 430

Asn Gly Ile Phe Asn Gly Leu Ser Asn Leu Glu Val Leu Lys Met Ala
435 440 445

Gly Asn Ser Phe Gln Glu Asn Phe Leu Pro Asp Ile Phe Thr Glu Leu
450 455 460

Arg Asn Leu Thr Phe Leu Asp Leu Ser Gln Cys Gln Leu Glu Gln Leu
465 470 475 480

Ser Pro Thr Ala Phe Asn Ser Leu Ser Ser Leu Gln Val Leu Asn Met
485 490 495

Ser His Asn Asn Phe Phe Ser Leu Asp Thr Phe Pro Tyr Lys Cys Leu
500 505 510

Asn Ser Leu Gln Val Leu Asp Tyr Ser Leu Asn His Ile Met Thr Ser
515 520 525

Lys Lys Gln Glu Leu Gln Arg Phe Pro Ser Ser Leu Ala Phe Leu Asn
530 535 540

Leu Thr Gln Asn Asp Phe Ala Cys Thr Cys Glu His Glu Ser Phe Leu
545 550 555 560

Gln Trp Ile Lys Asp Gln Arg Gln Leu Leu Val Glu Val Glu Arg Met
565 570 575

Glu Cys Ala Thr Pro Ser Asp Lys Gln Gly Met Pro Val Leu Ser Leu
580 585 590

Asn Ile Thr Cys Gln Met Asn Lys Thr Ile Ile Gly Val Ser Val Leu
595 600 605

Ser Val Leu Val Val Ser Val Val Ala Val Leu Val Tyr Lys Phe Tyr
610 615 620

Phe His Leu Met Leu Leu Ala Gly Cys Met Lys Tyr Gly Arg Gly Glu
625 630 635 640

Asn Thr Tyr Asp Ala Phe Val Ile Tyr Ser Ser Gln Asp Glu Asp Trp
645 650 655

Val Arg Asn Glu Leu Val Lys Asn Leu Glu Glu Gly Val Pro Pro Phe
660 665 670

Gln Leu Cys Leu His Tyr Arg Asp Phe Ile Xaa Gly Val Ala Ile Ala
675 680 685

Ala Asn Ile Ile His Glu Gly Phe His Lys Ser Arg Lys Val Ile Val
690 695 700

Val Val Ser Gln His Phe Ile Gln Ser Arg Trp Cys Ile Phe Glu Tyr
705 710 715 720

Glu Ile Ala Gln Thr Trp Gln Phe Leu Ser Ser His Ala Gly Ile Ile
725 730 735

Phe Ile Val Leu Gln Lys Val Glu Lys Thr Leu Leu Arg Gln Gln Val
740 745 750

Glu Leu Tyr Arg Leu Leu Ser Arg Asn Thr Tyr Leu Glu Trp Glu Asp
755 760 765

Ser Val Leu Gly Arg His Ile Phe Trp Arg Arg Leu Arg Lys Ala Leu
770 775 780

Leu Asp Gly Lys Ser Trp Asn Pro Glu Gly Thr Val Gly Thr Gly Cys
785 790 795 800

Asn

<210> 10
<211> 2388
<212> DNA
<213> Macaca mulatta

<400> 10
gtggttccta atattactta tcaatgcatg gagctgaatt tctacaaaat ccccgacaac 60

ctccccttct caaccaagaa cctggacctg agctttaatc ccctgaggca tttaggcagc 120

tatagcttct tcagtttccc agaactgcag gtgctggatt tatccaggtg tgaaatccag 180

acaattgaag atggggcata tcagagccta agccacctct ccactttaat attgacagga 240

aaccccatcc agagtttagc cctgggagcc ttttctggac tatcaagttt acagaagctg 300

gtggctgtgg agacaaatct agcatctcta gagaacttcc ccattggaca tctcaaaact 360

ttgaaagaac ttaatgtggc tcacaatctt atccagtctt tcaaattacc tgagtatttt 420

tctaatctga ccaatctaga gcacttggac ctttccagta acaagattca aaatatttat 480

tgcaaagact tgcaggttct acatcaaatg cccctatcca atctctcttt agacctgtcc 540

ctgaacccta taaactttat ccaaccaggt gcatttaaag aaattaggct tcataagctg 600

actttgagaa gtaattttga tgatttaaat gtaatgaaaa cttgtattca aggtctggct 660

ggtttagaag tccatcgttt ggttctggga gaatttagaa atgaaagaaa cttggaagag 720

tttgacaaat cttctctgga gggattgtgc aatttgacca ttgaagaatt ccgattaaca 780

tacttagact actacctcga taatattatt gacttattta attgtttggc aaatgtttct 840

tcattttccc tggtgagtgt gagtattaaa agggtagaag acttttctta taatttcaga 900

tggcaacatt tagaattagt taactgtaaa tttgaacagt ttcccacatt ggaactcgaa 960

tctctcaaaa ggcttacttt cactgccaac aaaggtggga atgctttttc agaagttgat 1020

ctaccaagcc ttgagtttct agatctcagt agaaatggct tgagtttcaa aggttgctgt 1080

tctcaaagtg attttgggac aaccagccta aagtatttag atctgagctt caatgatgtt 1140

attaccatga gttcaaactt cttgggctta gaaaaactag aacatctgga tttccagcat 1200

tccaatttga aacagatgag tcaattttca gtattcctat cactcagaaa cctcatttac 1260

cttgacattt ctcatactca caccagagtt gctttcaatg gcatcttcga tggcttgctc 1320

agtctcaaag tcttaaaaat ggctggcaat tctttccagg aaaacttcct tccagatatc 1380

ttcacagatc tgaaaaactt gaccttcctg gacctctctc agtgtcaatt ggagcagttg 1440

tctccaacag catttgacac actcaacaag cttcaggtac taaatatgag ccacaacaac 1500

ttcttttcat tggatacgtt tccttataag tgtctgccct ccctccaggt tctcgattac 1560

agtctcaatc acataatgac ttccaacaac caggaactac agcattttcc aagtagtcta 1620

gctttcttaa atcttactca gaatgacttt gcttgtactt gtgaacacca gagtttcctg 1680

cagtggatca aggaccagag gcagctcttg gtggaagctg aacgaatgga atgtgcaaca 1740

ccttcagata aacagggcat gccggtgctg agtttgaata ttacctgtca gatgaataag 1800

accatcattg gtgtgtctgt gttcagtgtg cttgtggtat ctgttgtagc agttctggtc 1860

tataagttct attttcacct gatgcttctt gctggctgca taaastatgg tagaggtgaa 1920

aacatctatg atgcctttgt tatctactca agccaggatg aggactgggt aaggaatgaa 1980

ctagtaaaga atttagaaga aggggtgcct ccctttcagc tctgccttca ctacagagac 2040

tttattcccg gtgtggccat tgctgcaaac atcatccatg aaggtttcca taaaagccga 2100

aaggtgattg ttgtggtgtc ccagcacttc atccagagcc gctggtgtat ctttgaatat 2160

gagattgctc agacctggca gtttctgagc agtcgtgcag gcataatctt cattgtcctg 2220

cagaaggtgg agaagaccct gctcaggcag caggtggagc tgtaccgcct tctcagcagg 2280

aacacttacc tggagtggga ggacagtgtc ctggggcagc acatcttctg gagacgactc 2340

agaaaagccc tgttggatgg cagatcgtgg aatccagaag aacagtag 2388

<210> 11
<211> 2388
<212> DNA
<213> Macaca mulatta

<220>
<221> CDS
<222> (1)..(2388)

<400> 11

```
gtg gtt cct aat att act tat caa tgc atg gag ctg aat ttc tac aaa       48
Val Val Pro Asn Ile Thr Tyr Gln Cys Met Glu Leu Asn Phe Tyr Lys
1               5                  10                  15

atc ccc gac aac ctc ccc ttc tca acc aag aac ctg gac ctg agc ttt       96
Ile Pro Asp Asn Leu Pro Phe Ser Thr Lys Asn Leu Asp Leu Ser Phe
            20                  25                  30

aat ccc ctg agg cat tta ggc agc tat agc ttc ttc agt ttc cca gaa      144
Asn Pro Leu Arg His Leu Gly Ser Tyr Ser Phe Phe Ser Phe Pro Glu
        35                  40                  45

ctg cag gtg ctg gat tta tcc agg tgt gaa atc cag aca att gaa gat      192
Leu Gln Val Leu Asp Leu Ser Arg Cys Glu Ile Gln Thr Ile Glu Asp
    50                  55                  60

ggg gca tat cag agc cta agc cac ctc tcc act tta ata ttg aca gga      240
Gly Ala Tyr Gln Ser Leu Ser His Leu Ser Thr Leu Ile Leu Thr Gly
65                  70                  75                  80

aac ccc atc cag agt tta gcc ctg gga gcc ttt tct gga cta tca agt      288
Asn Pro Ile Gln Ser Leu Ala Leu Gly Ala Phe Ser Gly Leu Ser Ser
                85                  90                  95

tta cag aag ctg gtg gct gtg gag aca aat cta gca tct cta gag aac      336
Leu Gln Lys Leu Val Ala Val Glu Thr Asn Leu Ala Ser Leu Glu Asn
            100                 105                 110

ttc ccc att gga cat ctc aaa act ttg aaa gaa ctt aat gtg gct cac      384
Phe Pro Ile Gly His Leu Lys Thr Leu Lys Glu Leu Asn Val Ala His
        115                 120                 125

aat ctt atc cag tct ttc aaa tta cct gag tat ttt tct aat ctg acc      432
Asn Leu Ile Gln Ser Phe Lys Leu Pro Glu Tyr Phe Ser Asn Leu Thr
    130                 135                 140

aat cta gag cac ttg gac ctt tcc agt aac aag att caa aat att tat      480
Asn Leu Glu His Leu Asp Leu Ser Ser Asn Lys Ile Gln Asn Ile Tyr
145                 150                 155                 160

tgc aaa gac ttg cag gtt cta cat caa atg ccc cta tcc aat ctc tct      528
Cys Lys Asp Leu Gln Val Leu His Gln Met Pro Leu Ser Asn Leu Ser
                165                 170                 175

tta gac ctg tcc ctg aac cct ata aac ttt atc caa cca ggt gca ttt      576
Leu Asp Leu Ser Leu Asn Pro Ile Asn Phe Ile Gln Pro Gly Ala Phe
            180                 185                 190

aaa gaa att agg ctt cat aag ctg act ttg aga agt aat ttt gat gat      624
Lys Glu Ile Arg Leu His Lys Leu Thr Leu Arg Ser Asn Phe Asp Asp
        195                 200                 205

tta aat gta atg aaa act tgt att caa ggt ctg gct ggt tta gaa gtc      672
Leu Asn Val Met Lys Thr Cys Ile Gln Gly Leu Ala Gly Leu Glu Val
    210                 215                 220

cat cgt ttg gtt ctg gga gaa ttt aga aat gaa aga aac ttg gaa gag      720
His Arg Leu Val Leu Gly Glu Phe Arg Asn Glu Arg Asn Leu Glu Glu
225                 230                 235                 240
```

```
ttt gac aaa tct tct ctg gag gga ttg tgc aat ttg acc att gaa gaa      768
Phe Asp Lys Ser Ser Leu Glu Gly Leu Cys Asn Leu Thr Ile Glu Glu
                245                 250                 255

ttc cga tta aca tac tta gac tac tac ctc gat aat att att gac tta      816
Phe Arg Leu Thr Tyr Leu Asp Tyr Tyr Leu Asp Asn Ile Ile Asp Leu
            260                 265                 270

ttt aat tgt ttg gca aat gtt tct tca ttt tcc ctg gtg agt gtg agt      864
Phe Asn Cys Leu Ala Asn Val Ser Ser Phe Ser Leu Val Ser Val Ser
        275                 280                 285

att aaa agg gta gaa gac ttt tct tat aat ttc aga tgg caa cat tta      912
Ile Lys Arg Val Glu Asp Phe Ser Tyr Asn Phe Arg Trp Gln His Leu
    290                 295                 300

gaa tta gtt aac tgt aaa ttt gaa cag ttt ccc aca ttg gaa ctc gaa      960
Glu Leu Val Asn Cys Lys Phe Glu Gln Phe Pro Thr Leu Glu Leu Glu
305                 310                 315                 320

tct ctc aaa agg ctt act ttc act gcc aac aaa ggt ggg aat gct ttt     1008
Ser Leu Lys Arg Leu Thr Phe Thr Ala Asn Lys Gly Gly Asn Ala Phe
                325                 330                 335

tca gaa gtt gat cta cca agc ctt gag ttt cta gat ctc agt aga aat     1056
Ser Glu Val Asp Leu Pro Ser Leu Glu Phe Leu Asp Leu Ser Arg Asn
            340                 345                 350

ggc ttg agt ttc aaa ggt tgc tgt tct caa agt gat ttt ggg aca acc     1104
Gly Leu Ser Phe Lys Gly Cys Cys Ser Gln Ser Asp Phe Gly Thr Thr
        355                 360                 365

agc cta aag tat tta gat ctg agc ttc aat gat gtt att acc atg agt     1152
Ser Leu Lys Tyr Leu Asp Leu Ser Phe Asn Asp Val Ile Thr Met Ser
    370                 375                 380

tca aac ttc ttg ggc tta gaa aaa cta gaa cat ctg gat ttc cag cat     1200
Ser Asn Phe Leu Gly Leu Glu Lys Leu Glu His Leu Asp Phe Gln His
385                 390                 395                 400

tcc aat ttg aaa cag atg agt caa ttt tca gta ttc cta tca ctc aga     1248
Ser Asn Leu Lys Gln Met Ser Gln Phe Ser Val Phe Leu Ser Leu Arg
                405                 410                 415

aac ctc att tac ctt gac att tct cat act cac acc aga gtt gct ttc     1296
Asn Leu Ile Tyr Leu Asp Ile Ser His Thr His Thr Arg Val Ala Phe
            420                 425                 430

aat ggc atc ttc gat ggc ttg ctc agt ctc aaa gtc tta aaa atg gct     1344
Asn Gly Ile Phe Asp Gly Leu Leu Ser Leu Lys Val Leu Lys Met Ala
        435                 440                 445

ggc aat tct ttc cag gaa aac ttc ctt cca gat atc ttc aca gat ctg     1392
Gly Asn Ser Phe Gln Glu Asn Phe Leu Pro Asp Ile Phe Thr Asp Leu
    450                 455                 460

aaa aac ttg acc ttc ctg gac ctc tct cag tgt caa ttg gag cag ttg     1440
Lys Asn Leu Thr Phe Leu Asp Leu Ser Gln Cys Gln Leu Glu Gln Leu
465                 470                 475                 480
```

```
tct cca aca gca ttt gac aca ctc aac aag ctt cag gta cta aat atg      1488
Ser Pro Thr Ala Phe Asp Thr Leu Asn Lys Leu Gln Val Leu Asn Met
                485                 490                 495

agc cac aac aac ttc ttt tca ttg gat acg ttt cct tat aag tgt ctg      1536
Ser His Asn Asn Phe Phe Ser Leu Asp Thr Phe Pro Tyr Lys Cys Leu
            500                 505                 510

ccc tcc ctc cag gtt ctc gat tac agt ctc aat cac ata atg act tcc      1584
Pro Ser Leu Gln Val Leu Asp Tyr Ser Leu Asn His Ile Met Thr Ser
        515                 520                 525

aac aac cag gaa cta cag cat ttt cca agt agt cta gct ttc tta aat      1632
Asn Asn Gln Glu Leu Gln His Phe Pro Ser Ser Leu Ala Phe Leu Asn
    530                 535                 540

ctt act cag aat gac ttt gct tgt act tgt gaa cac cag agt ttc ctg      1680
Leu Thr Gln Asn Asp Phe Ala Cys Thr Cys Glu His Gln Ser Phe Leu
545                 550                 555                 560

cag tgg atc aag gac cag agg cag ctc ttg gtg gaa gct gaa cga atg      1728
Gln Trp Ile Lys Asp Gln Arg Gln Leu Leu Val Glu Ala Glu Arg Met
                565                 570                 575

gaa tgt gca aca cct tca gat aaa cag ggc atg ccg gtg ctg agt ttg      1776
Glu Cys Ala Thr Pro Ser Asp Lys Gln Gly Met Pro Val Leu Ser Leu
            580                 585                 590

aat att acc tgt cag atg aat aag acc atc att ggt gtg tct gtg ttc      1824
Asn Ile Thr Cys Gln Met Asn Lys Thr Ile Ile Gly Val Ser Val Phe
        595                 600                 605

agt gtg ctt gtg gta tct gtt gta gca gtt ctg gtc tat aag ttc tat      1872
Ser Val Leu Val Val Ser Val Val Ala Val Leu Val Tyr Lys Phe Tyr
    610                 615                 620

ttt cac ctg atg ctt ctt gct ggc tgc ata aas tat ggt aga ggt gaa      1920
Phe His Leu Met Leu Leu Ala Gly Cys Ile Xaa Tyr Gly Arg Gly Glu
625                 630                 635                 640

aac atc tat gat gcc ttt gtt atc tac tca agc cag gat gag gac tgg      1968
Asn Ile Tyr Asp Ala Phe Val Ile Tyr Ser Ser Gln Asp Glu Asp Trp
                645                 650                 655

gta agg aat gaa cta gta aag aat tta gaa gaa ggg gtg cct ccc ttt      2016
Val Arg Asn Glu Leu Val Lys Asn Leu Glu Glu Gly Val Pro Pro Phe
            660                 665                 670

cag ctc tgc ctt cac tac aga gac ttt att ccc ggt gtg gcc att gct      2064
Gln Leu Cys Leu His Tyr Arg Asp Phe Ile Pro Gly Val Ala Ile Ala
        675                 680                 685

gca aac atc atc cat gaa ggt ttc cat aaa agc cga aag gtg att gtt      2112
Ala Asn Ile Ile His Glu Gly Phe His Lys Ser Arg Lys Val Ile Val
    690                 695                 700

gtg gtg tcc cag cac ttc atc cag agc cgc tgg tgt atc ttt gaa tat      2160
Val Val Ser Gln His Phe Ile Gln Ser Arg Trp Cys Ile Phe Glu Tyr
705                 710                 715                 720
```

```
gag att gct cag acc tgg cag ttt ctg agc agt cgt gca ggc ata atc     2208
Glu Ile Ala Gln Thr Trp Gln Phe Leu Ser Ser Arg Ala Gly Ile Ile
                725                 730                 735

ttc att gtc ctg cag aag gtg gag aag acc ctg ctc agg cag cag gtg     2256
Phe Ile Val Leu Gln Lys Val Glu Lys Thr Leu Leu Arg Gln Gln Val
            740                 745                 750

gag ctg tac cgc ctt ctc agc agg aac act tac ctg gag tgg gag gac     2304
Glu Leu Tyr Arg Leu Leu Ser Arg Asn Thr Tyr Leu Glu Trp Glu Asp
        755                 760                 765

agt gtc ctg ggg cag cac atc ttc tgg aga cga ctc aga aaa gcc ctg     2352
Ser Val Leu Gly Gln His Ile Phe Trp Arg Arg Leu Arg Lys Ala Leu
    770                 775                 780

ttg gat ggc aga tcg tgg aat cca gaa gaa cag tag                     2388
Leu Asp Gly Arg Ser Trp Asn Pro Glu Glu Gln
785                 790                 795
```

<210> 12
<211> 795
<212> PRT
<213> Macaca mulatta

<220>
<221> misc_feature
<222> (635)..(635)
<223> The 'Xaa' at location 635 stands for Lys, or Asn.

<400> 12

```
Val Val Pro Asn Ile Thr Tyr Gln Cys Met Glu Leu Asn Phe Tyr Lys
1               5                   10                  15

Ile Pro Asp Asn Leu Pro Phe Ser Thr Lys Asn Leu Asp Leu Ser Phe
            20                  25                  30

Asn Pro Leu Arg His Leu Gly Ser Tyr Ser Phe Phe Ser Phe Pro Glu
        35                  40                  45

Leu Gln Val Leu Asp Leu Ser Arg Cys Glu Ile Gln Thr Ile Glu Asp
    50                  55                  60

Gly Ala Tyr Gln Ser Leu Ser His Leu Ser Thr Leu Ile Leu Thr Gly
65                  70                  75                  80

Asn Pro Ile Gln Ser Leu Ala Leu Gly Ala Phe Ser Gly Leu Ser Ser
                85                  90                  95

Leu Gln Lys Leu Val Ala Val Glu Thr Asn Leu Ala Ser Leu Glu Asn
            100                 105                 110
```

Phe Pro Ile Gly His Leu Lys Thr Leu Lys Glu Leu Asn Val Ala His
115 120 125

Asn Leu Ile Gln Ser Phe Lys Leu Pro Glu Tyr Phe Ser Asn Leu Thr
130 135 140

Asn Leu Glu His Leu Asp Leu Ser Ser Asn Lys Ile Gln Asn Ile Tyr
145 150 155 160

Cys Lys Asp Leu Gln Val Leu His Gln Met Pro Leu Ser Asn Leu Ser
165 170 175

Leu Asp Leu Ser Leu Asn Pro Ile Asn Phe Ile Gln Pro Gly Ala Phe
180 185 190

Lys Glu Ile Arg Leu His Lys Leu Thr Leu Arg Ser Asn Phe Asp Asp
195 200 205

Leu Asn Val Met Lys Thr Cys Ile Gln Gly Leu Ala Gly Leu Glu Val
210 215 220

His Arg Leu Val Leu Gly Glu Phe Arg Asn Glu Arg Asn Leu Glu Glu
225 230 235 240

Phe Asp Lys Ser Ser Leu Glu Gly Leu Cys Asn Leu Thr Ile Glu Glu
245 250 255

Phe Arg Leu Thr Tyr Leu Asp Tyr Tyr Leu Asp Asn Ile Ile Asp Leu
260 265 270

Phe Asn Cys Leu Ala Asn Val Ser Ser Phe Ser Leu Val Ser Val Ser
275 280 285

Ile Lys Arg Val Glu Asp Phe Ser Tyr Asn Phe Arg Trp Gln His Leu
290 295 300

Glu Leu Val Asn Cys Lys Phe Glu Gln Phe Pro Thr Leu Glu Leu Glu
305 310 315 320

Ser Leu Lys Arg Leu Thr Phe Thr Ala Asn Lys Gly Gly Asn Ala Phe
325 330 335

Ser Glu Val Asp Leu Pro Ser Leu Glu Phe Leu Asp Leu Ser Arg Asn
340 345 350

```
Gly Leu Ser Phe Lys Gly Cys Cys Ser Gln Ser Asp Phe Gly Thr Thr
        355                 360                 365

Ser Leu Lys Tyr Leu Asp Leu Ser Phe Asn Asp Val Ile Thr Met Ser
    370                 375                 380

Ser Asn Phe Leu Gly Leu Glu Lys Leu Glu His Leu Asp Phe Gln His
385                 390                 395                 400

Ser Asn Leu Lys Gln Met Ser Gln Phe Ser Val Phe Leu Ser Leu Arg
                405                 410                 415

Asn Leu Ile Tyr Leu Asp Ile Ser His Thr His Thr Arg Val Ala Phe
            420                 425                 430

Asn Gly Ile Phe Asp Gly Leu Leu Ser Leu Lys Val Leu Lys Met Ala
        435                 440                 445

Gly Asn Ser Phe Gln Glu Asn Phe Leu Pro Asp Ile Phe Thr Asp Leu
    450                 455                 460

Lys Asn Leu Thr Phe Leu Asp Leu Ser Gln Cys Gln Leu Glu Gln Leu
465                 470                 475                 480

Ser Pro Thr Ala Phe Asp Thr Leu Asn Lys Leu Gln Val Leu Asn Met
                485                 490                 495

Ser His Asn Asn Phe Phe Ser Leu Asp Thr Phe Pro Tyr Lys Cys Leu
            500                 505                 510

Pro Ser Leu Gln Val Leu Asp Tyr Ser Leu Asn His Ile Met Thr Ser
        515                 520                 525

Asn Asn Gln Glu Leu Gln His Phe Pro Ser Ser Leu Ala Phe Leu Asn
    530                 535                 540

Leu Thr Gln Asn Asp Phe Ala Cys Thr Cys Glu His Gln Ser Phe Leu
545                 550                 555                 560

Gln Trp Ile Lys Asp Gln Arg Gln Leu Leu Val Glu Ala Glu Arg Met
                565                 570                 575

Glu Cys Ala Thr Pro Ser Asp Lys Gln Gly Met Pro Val Leu Ser Leu
            580                 585                 590
```

Asn Ile Thr Cys Gln Met Asn Lys Thr Ile Ile Gly Val Ser Val Phe
595 600 605

Ser Val Leu Val Val Ser Val Val Ala Val Leu Val Tyr Lys Phe Tyr
610 615 620

Phe His Leu Met Leu Leu Ala Gly Cys Ile Xaa Tyr Gly Arg Gly Glu
625 630 635 640

Asn Ile Tyr Asp Ala Phe Val Ile Tyr Ser Ser Gln Asp Glu Asp Trp
645 650 655

Val Arg Asn Glu Leu Val Lys Asn Leu Glu Glu Gly Val Pro Pro Phe
660 665 670

Gln Leu Cys Leu His Tyr Arg Asp Phe Ile Pro Gly Val Ala Ile Ala
675 680 685

Ala Asn Ile Ile His Glu Gly Phe His Lys Ser Arg Lys Val Ile Val
690 695 700

Val Val Ser Gln His Phe Ile Gln Ser Arg Trp Cys Ile Phe Glu Tyr
705 710 715 720

Glu Ile Ala Gln Thr Trp Gln Phe Leu Ser Ser Arg Ala Gly Ile Ile
725 730 735

Phe Ile Val Leu Gln Lys Val Glu Lys Thr Leu Leu Arg Gln Gln Val
740 745 750

Glu Leu Tyr Arg Leu Leu Ser Arg Asn Thr Tyr Leu Glu Trp Glu Asp
755 760 765

Ser Val Leu Gly Gln His Ile Phe Trp Arg Arg Leu Arg Lys Ala Leu
770 775 780

Leu Asp Gly Arg Ser Trp Asn Pro Glu Glu Gln
785 790 795

<210> 13
<211> 2238
<212> DNA
<213> Cebus capucinus

<400> 13

tgtgaaatcc acacaattga agatggtgca tatcagagcc taagccacct ctccacctta 60

atattgacag gaaatcctat ccagaattta gccctgggag ccttttctgg actatcaagt 120

ttacagaaac tggtagctgt ggagacacat ctgttatcgc tagaaagctt ccccattgga 180

catctcaaaa ctttgaagga ccttaatgtg gctcacaatc taatccaatc tttcaaatta 240

cctgagtatt tttctaatct gaccaatcta gagcacttgg acctttctag taacaatatt 300

caaaatattt attgcaaaga cttgcaggtt ctacatcaaa tgcccctact caatctctct 360

ttagacctgt ccctgaaccc tataaacttt attcagccag gtgcatttaa agaaattagg 420

ctccgtaagc tgactttgag aaataatttt gatagtttaa atgtaatgaa aacttgcatt 480

cacggtctgg ctgggttaga agtccatcgt ttggttctgg gagaatttag aaatgaaaga 540

aatattgaag actttgacaa atctgctctg gagggcctgt gcaatttgac catcaaagaa 600

ttccgattag catacttaga caactttcca gatgatatta ttgacttatt taattgtttg 660

gtaaatgttt cttcattttc cctgttgagt gtgtatatta aaagagtaga agacttttct 720

tataatttca gatggcaaca tttagaatta gttaactgta tatttcaaca gtttcctcca 780

ctgaaactca aatctctcaa aaggcttact ttcagtaaaa acaaaggtag gaatcatttt 840

gcagaagttg atctgccaag ccttgagttt ctagatctca gtagaaatgg cttgagtttc 900

aaaggttgct gttctcaatc tgattttggg acgaccagcc taaagtattt agatctgagc 960

ttcaatgatg ttattaccat gagttcaaac ttcttaggct tagaacaact agaacacttg 1020

gatttccagc attccaattt gaaacaaatg agtgagtttt cagtatttct atcactcaga 1080

aacctcattt accttgacat ttctcatact cacaccagag ttgctttcaa tggcatcttt 1140

aatggcttgt tcagtctcaa agtcttgaaa atggctggaa attctttcca gcaaaacttc 1200

cttgcagata tcttcacaga tctgaataac ttgatattcc tggacctttc tgagtgtcaa 1260

ctggagcagt tgtctccaac agcatttgac tcacttccca gacttcagat actaaatatg 1320

agccacaaca agttctttgc attggataca tttccttata agcatctcta ctccctccac 1380

gttctggatt acagtctcaa tcacataggg acttccaaaa atcaggaact acagcatttt 1440

ccaagtagtc tagctttctt aaatcttact caaaatgact ttgcttgtac ttgtgaacac 1500

cagagtttcc tgcagtggat caaggaccag aggcggctat tggtggaagt tgaacgaatg 1560

gaatgcgcaa cacctttaaa taggaagggc atacctgtgc tgagtttgaa tatcacctgt 1620

cagatgagta agaccatcat tggtgtgtca gtgctcagtg tgcttgtggt atctgttgta 1680

gcagttctgg tctataagtt ctattttcac ctgatgcttc ttgctggctg cataaagtat 1740

ggtagaggtg aaaacaccta tgatgccttt gttatctact caagccagga tgaggactgg 1800

gtaaggaatg aactagtaaa gaatttagaa gaaggggtgc ctccttttca gctctgcctt 1860

cactacagag actttattcc cggtgtggcc attgctgcca acatcatcca tgaaggtttc 1920

cataaaagcc gaaaggtgat tgttgtggta tcccagcact tcatccagag ccgctggtgt 1980

atctttgaat atgagattgc tcagacctgg cagtttctga gcagtcgtgc tggtatcatc 2040

ttcattgtcc tgcagaaggt ggagaagtcc ctgctcaggc agcaggtgga gctgtaccgc 2100

cttctcagca ggaacaccta cctggagtgg gaggacagtg tcctggggag gcatatcttc 2160

tggaggcgac tcagaaaagc cctgctgaat ggtagaccgt ggagtccaga aggaacagtg 2220

ggtgcaggat gcgattag 2238

<210> 14
<211> 2238
<212> DNA
<213> Cebus capucinus

<220>
<221> CDS
<222> (1)..(2238)

<400> 14

```
tgt gaa atc cac aca att gaa gat ggt gca tat cag agc cta agc cac      48
Cys Glu Ile His Thr Ile Glu Asp Gly Ala Tyr Gln Ser Leu Ser His
1               5                  10                  15

ctc tcc acc tta ata ttg aca gga aat cct atc cag aat tta gcc ctg      96
Leu Ser Thr Leu Ile Leu Thr Gly Asn Pro Ile Gln Asn Leu Ala Leu
            20                  25                  30

gga gcc ttt tct gga cta tca agt tta cag aaa ctg gta gct gtg gag     144
Gly Ala Phe Ser Gly Leu Ser Ser Leu Gln Lys Leu Val Ala Val Glu
        35                  40                  45

aca cat ctg tta tcg cta gaa agc ttc ccc att gga cat ctc aaa act     192
Thr His Leu Leu Ser Leu Glu Ser Phe Pro Ile Gly His Leu Lys Thr
    50                  55                  60

ttg aag gac ctt aat gtg gct cac aat cta atc caa tct ttc aaa tta     240
Leu Lys Asp Leu Asn Val Ala His Asn Leu Ile Gln Ser Phe Lys Leu
65                  70                  75                  80

cct gag tat ttt tct aat ctg acc aat cta gag cac ttg gac ctt tct     288
Pro Glu Tyr Phe Ser Asn Leu Thr Asn Leu Glu His Leu Asp Leu Ser
                85                  90                  95

agt aac aat att caa aat att tat tgc aaa gac ttg cag gtt cta cat     336
Ser Asn Asn Ile Gln Asn Ile Tyr Cys Lys Asp Leu Gln Val Leu His
            100                 105                 110

caa atg ccc cta ctc aat ctc tct tta gac ctg tcc ctg aac cct ata     384
Gln Met Pro Leu Leu Asn Leu Ser Leu Asp Leu Ser Leu Asn Pro Ile
        115                 120                 125
```

```
aac ttt att cag cca ggt gca ttt aaa gaa att agg ctc cgt aag ctg      432
Asn Phe Ile Gln Pro Gly Ala Phe Lys Glu Ile Arg Leu Arg Lys Leu
    130                 135                 140

act ttg aga aat aat ttt gat agt tta aat gta atg aaa act tgc att      480
Thr Leu Arg Asn Asn Phe Asp Ser Leu Asn Val Met Lys Thr Cys Ile
145                 150                 155                 160

cac ggt ctg gct ggg tta gaa gtc cat cgt ttg gtt ctg gga gaa ttt      528
His Gly Leu Ala Gly Leu Glu Val His Arg Leu Val Leu Gly Glu Phe
                165                 170                 175

aga aat gaa aga aat att gaa gac ttt gac aaa tct gct ctg gag ggc      576
Arg Asn Glu Arg Asn Ile Glu Asp Phe Asp Lys Ser Ala Leu Glu Gly
            180                 185                 190

ctg tgc aat ttg acc atc aaa gaa ttc cga tta gca tac tta gac aac      624
Leu Cys Asn Leu Thr Ile Lys Glu Phe Arg Leu Ala Tyr Leu Asp Asn
        195                 200                 205

ttt cca gat gat att att gac tta ttt aat tgt ttg gta aat gtt tct      672
Phe Pro Asp Asp Ile Ile Asp Leu Phe Asn Cys Leu Val Asn Val Ser
    210                 215                 220

tca ttt tcc ctg ttg agt gtg tat att aaa aga gta gaa gac ttt tct      720
Ser Phe Ser Leu Leu Ser Val Tyr Ile Lys Arg Val Glu Asp Phe Ser
225                 230                 235                 240

tat aat ttc aga tgg caa cat tta gaa tta gtt aac tgt ata ttt caa      768
Tyr Asn Phe Arg Trp Gln His Leu Glu Leu Val Asn Cys Ile Phe Gln
                245                 250                 255

cag ttt cct cca ctg aaa ctc aaa tct ctc aaa agg ctt act ttc agt      816
Gln Phe Pro Pro Leu Lys Leu Lys Ser Leu Lys Arg Leu Thr Phe Ser
            260                 265                 270

aaa aac aaa ggt agg aat cat ttt gca gaa gtt gat ctg cca agc ctt      864
Lys Asn Lys Gly Arg Asn His Phe Ala Glu Val Asp Leu Pro Ser Leu
        275                 280                 285

gag ttt cta gat ctc agt aga aat ggc ttg agt ttc aaa ggt tgc tgt      912
Glu Phe Leu Asp Leu Ser Arg Asn Gly Leu Ser Phe Lys Gly Cys Cys
    290                 295                 300

tct caa tct gat ttt ggg acg acc agc cta aag tat tta gat ctg agc      960
Ser Gln Ser Asp Phe Gly Thr Thr Ser Leu Lys Tyr Leu Asp Leu Ser
305                 310                 315                 320

ttc aat gat gtt att acc atg agt tca aac ttc tta ggc tta gaa caa     1008
Phe Asn Asp Val Ile Thr Met Ser Ser Asn Phe Leu Gly Leu Glu Gln
                325                 330                 335

cta gaa cac ttg gat ttc cag cat tcc aat ttg aaa caa atg agt gag     1056
Leu Glu His Leu Asp Phe Gln His Ser Asn Leu Lys Gln Met Ser Glu
            340                 345                 350

ttt tca gta ttt cta tca ctc aga aac ctc att tac ctt gac att tct     1104
Phe Ser Val Phe Leu Ser Leu Arg Asn Leu Ile Tyr Leu Asp Ile Ser
        355                 360                 365
```

```
cat act cac acc aga gtt gct ttc aat ggc atc ttt aat ggc ttg ttc      1152
His Thr His Thr Arg Val Ala Phe Asn Gly Ile Phe Asn Gly Leu Phe
    370                 375                 380

agt ctc aaa gtc ttg aaa atg gct gga aat tct ttc cag caa aac ttc      1200
Ser Leu Lys Val Leu Lys Met Ala Gly Asn Ser Phe Gln Gln Asn Phe
385                 390                 395                 400

ctt gca gat atc ttc aca gat ctg aat aac ttg ata ttc ctg gac ctt      1248
Leu Ala Asp Ile Phe Thr Asp Leu Asn Asn Leu Ile Phe Leu Asp Leu
                405                 410                 415

tct gag tgt caa ctg gag cag ttg tct cca aca gca ttt gac tca ctt      1296
Ser Glu Cys Gln Leu Glu Gln Leu Ser Pro Thr Ala Phe Asp Ser Leu
            420                 425                 430

ccc aga ctt cag ata cta aat atg agc cac aac aag ttc ttt gca ttg      1344
Pro Arg Leu Gln Ile Leu Asn Met Ser His Asn Lys Phe Phe Ala Leu
        435                 440                 445

gat aca ttt cct tat aag cat ctc tac tcc ctc cac gtt ctg gat tac      1392
Asp Thr Phe Pro Tyr Lys His Leu Tyr Ser Leu His Val Leu Asp Tyr
    450                 455                 460

agt ctc aat cac ata ggg act tcc aaa aat cag gaa cta cag cat ttt      1440
Ser Leu Asn His Ile Gly Thr Ser Lys Asn Gln Glu Leu Gln His Phe
465                 470                 475                 480

cca agt agt cta gct ttc tta aat ctt act caa aat gac ttt gct tgt      1488
Pro Ser Ser Leu Ala Phe Leu Asn Leu Thr Gln Asn Asp Phe Ala Cys
                485                 490                 495

act tgt gaa cac cag agt ttc ctg cag tgg atc aag gac cag agg cgg      1536
Thr Cys Glu His Gln Ser Phe Leu Gln Trp Ile Lys Asp Gln Arg Arg
            500                 505                 510

cta ttg gtg gaa gtt gaa cga atg gaa tgc gca aca cct tta aat agg      1584
Leu Leu Val Glu Val Glu Arg Met Glu Cys Ala Thr Pro Leu Asn Arg
        515                 520                 525

aag ggc ata cct gtg ctg agt ttg aat atc acc tgt cag atg agt aag      1632
Lys Gly Ile Pro Val Leu Ser Leu Asn Ile Thr Cys Gln Met Ser Lys
    530                 535                 540

acc atc att ggt gtg tca gtg ctc agt gtg ctt gtg gta tct gtt gta      1680
Thr Ile Ile Gly Val Ser Val Leu Ser Val Leu Val Val Ser Val Val
545                 550                 555                 560

gca gtt ctg gtc tat aag ttc tat ttt cac ctg atg ctt ctt gct ggc      1728
Ala Val Leu Val Tyr Lys Phe Tyr Phe His Leu Met Leu Leu Ala Gly
                565                 570                 575

tgc ata aag tat ggt aga ggt gaa aac acc tat gat gcc ttt gtt atc      1776
Cys Ile Lys Tyr Gly Arg Gly Glu Asn Thr Tyr Asp Ala Phe Val Ile
            580                 585                 590

tac tca agc cag gat gag gac tgg gta agg aat gaa cta gta aag aat      1824
Tyr Ser Ser Gln Asp Glu Asp Trp Val Arg Asn Glu Leu Val Lys Asn
        595                 600                 605
```

```
tta gaa gaa ggg gtg cct cct ttt cag ctc tgc ctt cac tac aga gac     1872
Leu Glu Glu Gly Val Pro Pro Phe Gln Leu Cys Leu His Tyr Arg Asp
    610                 615                 620

ttt att ccc ggt gtg gcc att gct gcc aac atc atc cat gaa ggt ttc     1920
Phe Ile Pro Gly Val Ala Ile Ala Ala Asn Ile Ile His Glu Gly Phe
625                 630                 635                 640

cat aaa agc cga aag gtg att gtt gtg gta tcc cag cac ttc atc cag     1968
His Lys Ser Arg Lys Val Ile Val Val Val Ser Gln His Phe Ile Gln
                645                 650                 655

agc cgc tgg tgt atc ttt gaa tat gag att gct cag acc tgg cag ttt     2016
Ser Arg Trp Cys Ile Phe Glu Tyr Glu Ile Ala Gln Thr Trp Gln Phe
            660                 665                 670

ctg agc agt cgt gct ggt atc atc ttc att gtc ctg cag aag gtg gag     2064
Leu Ser Ser Arg Ala Gly Ile Ile Phe Ile Val Leu Gln Lys Val Glu
        675                 680                 685

aag tcc ctg ctc agg cag cag gtg gag ctg tac cgc ctt ctc agc agg     2112
Lys Ser Leu Leu Arg Gln Gln Val Glu Leu Tyr Arg Leu Leu Ser Arg
    690                 695                 700

aac acc tac ctg gag tgg gag gac agt gtc ctg ggg agg cat atc ttc     2160
Asn Thr Tyr Leu Glu Trp Glu Asp Ser Val Leu Gly Arg His Ile Phe
705                 710                 715                 720

tgg agg cga ctc aga aaa gcc ctg ctg aat ggt aga ccg tgg agt cca     2208
Trp Arg Arg Leu Arg Lys Ala Leu Leu Asn Gly Arg Pro Trp Ser Pro
                725                 730                 735

gaa gga aca gtg ggt gca gga tgc gat tag                             2238
Glu Gly Thr Val Gly Ala Gly Cys Asp
            740                 745
```

<210> 15
<211> 745
<212> PRT
<213> Cebus capucinus

<400> 15

```
Cys Glu Ile His Thr Ile Glu Asp Gly Ala Tyr Gln Ser Leu Ser His
1               5                   10                  15

Leu Ser Thr Leu Ile Leu Thr Gly Asn Pro Ile Gln Asn Leu Ala Leu
            20                  25                  30

Gly Ala Phe Ser Gly Leu Ser Ser Leu Gln Lys Leu Val Ala Val Glu
        35                  40                  45

Thr His Leu Leu Ser Leu Glu Ser Phe Pro Ile Gly His Leu Lys Thr
    50                  55                  60
```

Leu Lys Asp Leu Asn Val Ala His Asn Leu Ile Gln Ser Phe Lys Leu
65 70 75 80

Pro Glu Tyr Phe Ser Asn Leu Thr Asn Leu Glu His Leu Asp Leu Ser
85 90 95

Ser Asn Asn Ile Gln Asn Ile Tyr Cys Lys Asp Leu Gln Val Leu His
100 105 110

Gln Met Pro Leu Leu Asn Leu Ser Leu Asp Leu Ser Leu Asn Pro Ile
115 120 125

Asn Phe Ile Gln Pro Gly Ala Phe Lys Glu Ile Arg Leu Arg Lys Leu
130 135 140

Thr Leu Arg Asn Asn Phe Asp Ser Leu Asn Val Met Lys Thr Cys Ile
145 150 155 160

His Gly Leu Ala Gly Leu Glu Val His Arg Leu Val Leu Gly Glu Phe
165 170 175

Arg Asn Glu Arg Asn Ile Glu Asp Phe Asp Lys Ser Ala Leu Glu Gly
180 185 190

Leu Cys Asn Leu Thr Ile Lys Glu Phe Arg Leu Ala Tyr Leu Asp Asn
195 200 205

Phe Pro Asp Asp Ile Ile Asp Leu Phe Asn Cys Leu Val Asn Val Ser
210 215 220

Ser Phe Ser Leu Leu Ser Val Tyr Ile Lys Arg Val Glu Asp Phe Ser
225 230 235 240

Tyr Asn Phe Arg Trp Gln His Leu Glu Leu Val Asn Cys Ile Phe Gln
245 250 255

Gln Phe Pro Pro Leu Lys Leu Lys Ser Leu Lys Arg Leu Thr Phe Ser
260 265 270

Lys Asn Lys Gly Arg Asn His Phe Ala Glu Val Asp Leu Pro Ser Leu
275 280 285

Glu Phe Leu Asp Leu Ser Arg Asn Gly Leu Ser Phe Lys Gly Cys Cys
290 295 300

Ser Gln Ser Asp Phe Gly Thr Thr Ser Leu Lys Tyr Leu Asp Leu Ser
305 310 315 320

Phe Asn Asp Val Ile Thr Met Ser Ser Asn Phe Leu Gly Leu Glu Gln
325 330 335

Leu Glu His Leu Asp Phe Gln His Ser Asn Leu Lys Gln Met Ser Glu
340 345 350

Phe Ser Val Phe Leu Ser Leu Arg Asn Leu Ile Tyr Leu Asp Ile Ser
355 360 365

His Thr His Thr Arg Val Ala Phe Asn Gly Ile Phe Asn Gly Leu Phe
370 375 380

Ser Leu Lys Val Leu Lys Met Ala Gly Asn Ser Phe Gln Gln Asn Phe
385 390 395 400

Leu Ala Asp Ile Phe Thr Asp Leu Asn Asn Leu Ile Phe Leu Asp Leu
405 410 415

Ser Glu Cys Gln Leu Glu Gln Leu Ser Pro Thr Ala Phe Asp Ser Leu
420 425 430

Pro Arg Leu Gln Ile Leu Asn Met Ser His Asn Lys Phe Phe Ala Leu
435 440 445

Asp Thr Phe Pro Tyr Lys His Leu Tyr Ser Leu His Val Leu Asp Tyr
450 455 460

Ser Leu Asn His Ile Gly Thr Ser Lys Asn Gln Glu Leu Gln His Phe
465 470 475 480

Pro Ser Ser Leu Ala Phe Leu Asn Leu Thr Gln Asn Asp Phe Ala Cys
485 490 495

Thr Cys Glu His Gln Ser Phe Leu Gln Trp Ile Lys Asp Gln Arg Arg
500 505 510

Leu Leu Val Glu Val Glu Arg Met Glu Cys Ala Thr Pro Leu Asn Arg
515 520 525

Lys Gly Ile Pro Val Leu Ser Leu Asn Ile Thr Cys Gln Met Ser Lys
530 535 540

Thr Ile Ile Gly Val Ser Val Leu Ser Val Leu Val Val Ser Val Val
545 550 555 560

Ala Val Leu Val Tyr Lys Phe Tyr Phe His Leu Met Leu Leu Ala Gly
565 570 575

Cys Ile Lys Tyr Gly Arg Gly Glu Asn Thr Tyr Asp Ala Phe Val Ile
580 585 590

Tyr Ser Ser Gln Asp Glu Asp Trp Val Arg Asn Glu Leu Val Lys Asn
595 600 605

Leu Glu Glu Gly Val Pro Pro Phe Gln Leu Cys Leu His Tyr Arg Asp
610 615 620

Phe Ile Pro Gly Val Ala Ile Ala Ala Asn Ile Ile His Glu Gly Phe
625 630 635 640

His Lys Ser Arg Lys Val Ile Val Val Ser Gln His Phe Ile Gln
645 650 655

Ser Arg Trp Cys Ile Phe Glu Tyr Glu Ile Ala Gln Thr Trp Gln Phe
660 665 670

Leu Ser Ser Arg Ala Gly Ile Ile Phe Ile Val Leu Gln Lys Val Glu
675 680 685

Lys Ser Leu Leu Arg Gln Gln Val Glu Leu Tyr Arg Leu Leu Ser Arg
690 695 700

Asn Thr Tyr Leu Glu Trp Glu Asp Ser Val Leu Gly Arg His Ile Phe
705 710 715 720

Trp Arg Arg Leu Arg Lys Ala Leu Leu Asn Gly Arg Pro Trp Ser Pro
725 730 735

Glu Gly Thr Val Gly Ala Gly Cys Asp
740 745

<210> 16
<211> 2406
<212> DNA
<213> Saimiri sciureus

<400> 16

gtggttccta acgttactta tcaatgcatg gaactgaaty tctacaaaat ccccgacaac 60

atccccttct caactaagaa cctggacctg agctttaacc ccctgaggca tttaggcagc 120

catagcttct tcaatttccc agaactgcag gtgctggatt tatccaggtg tgacatccag 180

acaatcgaag atggggcata tcagagccta agccacctct ccaccttaat attgacagga 240

aatcctatcc agaatttagc cctgggagcc ttttctggac tatcaagttt acagaagctg 300

gtggctgtgg agacacatct gttatcacta gagaacttcc ccattggaca tctcaaaact 360

ttgaaggacc ttaatgtggc tcacaatcta atccaatctt tcaaattacc tgagtatttt 420

tctaatctga ccaatctaga gcacttggac ctttctagta acaatattca aaatatttat 480

tgcaaagact tgcaggttct acatcaaatg cccctactca atctctcttt agacctgtcc 540

ctgaacccta taaactttat tcaaccaggt gcgtttaaag aaattaggct ccataagctg 600

actttgagaa ataattttga tagtttaaat gcaatgaaaa cttgcattca aggtctggct 660

gggttagaag tccatcgttt ggttctggga gaatttagaa atgaaagaaa tattgaagac 720

tttgacaaat ctgctctgga gggcctgtgc aatttgacca ttaatgaatt ccgattagct 780

tacttagatg actttctaga tgatattatt gacttattta actgtttagc aaatgtttct 840

tcattttccc tggtgaatgt gcatattaaa agagtagaag acttttctta taattttaga 900

tggcaacatt tagaattagt taactgtgta tttcaacagt ttcctccact gaaactcaaa 960

tctctcaaaa ggcttacttt cactgccaac aaaggtagga atcatttttc agaagttgat 1020

cttccaagcc ttgagtttct agatctcagt agaaatggct tgagtttcaa aggttgctgt 1080

tctcaatctg attttgggac gaccagccta aagtatttag atctgagctt caatgacgtt 1140

attaccatgg gttcaaactt cttaggctta gaacaactag aacacttgga tttccagcat 1200

tccaatttga aacaaatgag tgagttttca gtattcctat cactcagaaa cctcatttac 1260

cttgacattt ctcatactca caccagagtt gctttcaatg gcatctttaa tggcttgttc 1320

agtctcaaag tcttgaaaat ggctggaaat tctttccagc aaaacttcct tgaagatatc 1380

ttcacrgatc tgaataactt gatattcctg gacctctctg agtgtcagct ggagcagttg 1440

tctccaacag catttgactc acttcccaga cttcggatac taaatatgag ccacaacaac 1500

ttctttgcat tggatacatt cccttacaag catctctact ccctccaggt tctggattac 1560

agtctcaatc atatagggac ttccaaaaat caggaactgc agcattttcc aagtagtcta 1620

gctttcttaa atcttactca aaatgacttt gcttgtactt gtgaacacca gagtttcctg 1680

cagtggatca aggaccagag gcggctgttg gtggaagttg aacaaatgga atgtgcaaca 1740

cctttaaata ggaagggcat acctgtgctg agtttgaata tcacctgtca gatgagtaag 1800

actatcattg gtgtgtcagt gctcagtgtg cttgtggtat ctgttgtagc agttctggtc 1860

tataagttct attttcacct gatgcttctt gctggctgca taaagtatgg tagaggtgaa 1920

aacacctatg atgcctttgt tatctactca agccaggatg aggactgggt aaggaatgaa 1980

ctagtaaaga atttagaaga aggggtgcct ccctttcagc tctgccttca ctacagagac 2040

tttattcccg gtgtggccat tgctgccaac atcatccatg aaggtttcca taaaagccga 2100

aaggtgattg ttgtggtatc tcagcacttc atccagagcc gctggtgtat ctttgaatat 2160

gagattgctc agacctggca gtttctgagc agtcgtgctg gtatcatctt cattgtcctg 2220

cagaaggtgg agaagtccct gctcaggcag caggtggagc tgtaccgcct tctcagcagg 2280

aacacttacc tggagtggga ggacagtgtc ctggggaggc acatcttctg gagacgactc 2340

agaaaagccc tgctggatgg tagaccgtgg aatccagaag gaacagtggg tgcaggatgc 2400

gaatag 2406

<210> 17
<211> 2406
<212> DNA
<213> Saimiri sciureus

<220>
<221> CDS
<222> (1)..(2406)

<400> 17

```
gtg gtt cct aac gtt act tat caa tgc atg gaa ctg aat ytc tac aaa       48
Val Val Pro Asn Val Thr Tyr Gln Cys Met Glu Leu Asn Xaa Tyr Lys
1               5                  10                  15

atc ccc gac aac atc ccc ttc tca act aag aac ctg gac ctg agc ttt       96
Ile Pro Asp Asn Ile Pro Phe Ser Thr Lys Asn Leu Asp Leu Ser Phe
            20                  25                  30

aac ccc ctg agg cat tta ggc agc cat agc ttc ttc aat ttc cca gaa      144
Asn Pro Leu Arg His Leu Gly Ser His Ser Phe Phe Asn Phe Pro Glu
        35                  40                  45

ctg cag gtg ctg gat tta tcc agg tgt gac atc cag aca atc gaa gat      192
Leu Gln Val Leu Asp Leu Ser Arg Cys Asp Ile Gln Thr Ile Glu Asp
    50                  55                  60

ggg gca tat cag agc cta agc cac ctc tcc acc tta ata ttg aca gga      240
Gly Ala Tyr Gln Ser Leu Ser His Leu Ser Thr Leu Ile Leu Thr Gly
65                  70                  75                  80

aat cct atc cag aat tta gcc ctg gga gcc ttt tct gga cta tca agt      288
Asn Pro Ile Gln Asn Leu Ala Leu Gly Ala Phe Ser Gly Leu Ser Ser
                85                  90                  95

tta cag aag ctg gtg gct gtg gag aca cat ctg tta tca cta gag aac      336
Leu Gln Lys Leu Val Ala Val Glu Thr His Leu Leu Ser Leu Glu Asn
```

```
                100                 105                 110

ttc ccc att gga cat ctc aaa act ttg aag gac ctt aat gtg gct cac      384
Phe Pro Ile Gly His Leu Lys Thr Leu Lys Asp Leu Asn Val Ala His
        115                 120                 125

aat cta atc caa tct ttc aaa tta cct gag tat ttt tct aat ctg acc      432
Asn Leu Ile Gln Ser Phe Lys Leu Pro Glu Tyr Phe Ser Asn Leu Thr
    130                 135                 140

aat cta gag cac ttg gac ctt tct agt aac aat att caa aat att tat      480
Asn Leu Glu His Leu Asp Leu Ser Ser Asn Asn Ile Gln Asn Ile Tyr
145                 150                 155                 160

tgc aaa gac ttg cag gtt cta cat caa atg ccc cta ctc aat ctc tct      528
Cys Lys Asp Leu Gln Val Leu His Gln Met Pro Leu Leu Asn Leu Ser
                165                 170                 175

tta gac ctg tcc ctg aac cct ata aac ttt att caa cca ggt gcg ttt      576
Leu Asp Leu Ser Leu Asn Pro Ile Asn Phe Ile Gln Pro Gly Ala Phe
            180                 185                 190

aaa gaa att agg ctc cat aag ctg act ttg aga aat aat ttt gat agt      624
Lys Glu Ile Arg Leu His Lys Leu Thr Leu Arg Asn Asn Phe Asp Ser
        195                 200                 205

tta aat gca atg aaa act tgc att caa ggt ctg gct ggg tta gaa gtc      672
Leu Asn Ala Met Lys Thr Cys Ile Gln Gly Leu Ala Gly Leu Glu Val
    210                 215                 220

cat cgt ttg gtt ctg gga gaa ttt aga aat gaa aga aat att gaa gac      720
His Arg Leu Val Leu Gly Glu Phe Arg Asn Glu Arg Asn Ile Glu Asp
225                 230                 235                 240

ttt gac aaa tct gct ctg gag ggc ctg tgc aat ttg acc att aat gaa      768
Phe Asp Lys Ser Ala Leu Glu Gly Leu Cys Asn Leu Thr Ile Asn Glu
                245                 250                 255

ttc cga tta gct tac tta gat gac ttt cta gat gat att att gac tta      816
Phe Arg Leu Ala Tyr Leu Asp Asp Phe Leu Asp Asp Ile Ile Asp Leu
            260                 265                 270

ttt aac tgt tta gca aat gtt tct tca ttt tcc ctg gtg aat gtg cat      864
Phe Asn Cys Leu Ala Asn Val Ser Ser Phe Ser Leu Val Asn Val His
        275                 280                 285

att aaa aga gta gaa gac ttt tct tat aat ttt aga tgg caa cat tta      912
Ile Lys Arg Val Glu Asp Phe Ser Tyr Asn Phe Arg Trp Gln His Leu
    290                 295                 300

gaa tta gtt aac tgt gta ttt caa cag ttt cct cca ctg aaa ctc aaa      960
Glu Leu Val Asn Cys Val Phe Gln Gln Phe Pro Pro Leu Lys Leu Lys
305                 310                 315                 320

tct ctc aaa agg ctt act ttc act gcc aac aaa ggt agg aat cat ttt     1008
Ser Leu Lys Arg Leu Thr Phe Thr Ala Asn Lys Gly Arg Asn His Phe
                325                 330                 335

tca gaa gtt gat ctt cca agc ctt gag ttt cta gat ctc agt aga aat     1056
Ser Glu Val Asp Leu Pro Ser Leu Glu Phe Leu Asp Leu Ser Arg Asn
```

```
            340                 345                 350

ggc ttg agt ttc aaa ggt tgc tgt tct caa tct gat ttt ggg acg acc     1104
Gly Leu Ser Phe Lys Gly Cys Cys Ser Gln Ser Asp Phe Gly Thr Thr
        355                 360                 365

agc cta aag tat tta gat ctg agc ttc aat gac gtt att acc atg ggt     1152
Ser Leu Lys Tyr Leu Asp Leu Ser Phe Asn Asp Val Ile Thr Met Gly
    370                 375                 380

tca aac ttc tta ggc tta gaa caa cta gaa cac ttg gat ttc cag cat     1200
Ser Asn Phe Leu Gly Leu Glu Gln Leu Glu His Leu Asp Phe Gln His
385                 390                 395                 400

tcc aat ttg aaa caa atg agt gag ttt tca gta ttc cta tca ctc aga     1248
Ser Asn Leu Lys Gln Met Ser Glu Phe Ser Val Phe Leu Ser Leu Arg
                405                 410                 415

aac ctc att tac ctt gac att tct cat act cac acc aga gtt gct ttc     1296
Asn Leu Ile Tyr Leu Asp Ile Ser His Thr His Thr Arg Val Ala Phe
            420                 425                 430

aat ggc atc ttt aat ggc ttg ttc agt ctc aaa gtc ttg aaa atg gct     1344
Asn Gly Ile Phe Asn Gly Leu Phe Ser Leu Lys Val Leu Lys Met Ala
        435                 440                 445

gga aat tct ttc cag caa aac ttc ctt gaa gat atc ttc acr gat ctg     1392
Gly Asn Ser Phe Gln Gln Asn Phe Leu Glu Asp Ile Phe Xaa Asp Leu
    450                 455                 460

aat aac ttg ata ttc ctg gac ctc tct gag tgt cag ctg gag cag ttg     1440
Asn Asn Leu Ile Phe Leu Asp Leu Ser Glu Cys Gln Leu Glu Gln Leu
465                 470                 475                 480

tct cca aca gca ttt gac tca ctt ccc aga ctt cgg ata cta aat atg     1488
Ser Pro Thr Ala Phe Asp Ser Leu Pro Arg Leu Arg Ile Leu Asn Met
                485                 490                 495

agc cac aac aac ttc ttt gca ttg gat aca ttc cct tac aag cat ctc     1536
Ser His Asn Asn Phe Phe Ala Leu Asp Thr Phe Pro Tyr Lys His Leu
            500                 505                 510

tac tcc ctc cag gtt ctg gat tac agt ctc aat cat ata ggg act tcc     1584
Tyr Ser Leu Gln Val Leu Asp Tyr Ser Leu Asn His Ile Gly Thr Ser
        515                 520                 525

aaa aat cag gaa ctg cag cat ttt cca agt agt cta gct ttc tta aat     1632
Lys Asn Gln Glu Leu Gln His Phe Pro Ser Ser Leu Ala Phe Leu Asn
    530                 535                 540

ctt act caa aat gac ttt gct tgt act tgt gaa cac cag agt ttc ctg     1680
Leu Thr Gln Asn Asp Phe Ala Cys Thr Cys Glu His Gln Ser Phe Leu
545                 550                 555                 560

cag tgg atc aag gac cag agg cgg ctg ttg gtg gaa gtt gaa caa atg     1728
Gln Trp Ile Lys Asp Gln Arg Arg Leu Leu Val Glu Val Glu Gln Met
                565                 570                 575

gaa tgt gca aca cct tta aat agg aag ggc ata cct gtg ctg agt ttg     1776
Glu Cys Ala Thr Pro Leu Asn Arg Lys Gly Ile Pro Val Leu Ser Leu
```

```
            580                 585                 590

aat atc acc tgt cag atg agt aag act atc att ggt gtg tca gtg ctc     1824
Asn Ile Thr Cys Gln Met Ser Lys Thr Ile Ile Gly Val Ser Val Leu
        595                 600                 605

agt gtg ctt gtg gta tct gtt gta gca gtt ctg gtc tat aag ttc tat     1872
Ser Val Leu Val Val Ser Val Val Ala Val Leu Val Tyr Lys Phe Tyr
    610                 615                 620

ttt cac ctg atg ctt ctt gct ggc tgc ata aag tat ggt aga ggt gaa     1920
Phe His Leu Met Leu Leu Ala Gly Cys Ile Lys Tyr Gly Arg Gly Glu
625                 630                 635                 640

aac acc tat gat gcc ttt gtt atc tac tca agc cag gat gag gac tgg     1968
Asn Thr Tyr Asp Ala Phe Val Ile Tyr Ser Ser Gln Asp Glu Asp Trp
                645                 650                 655

gta agg aat gaa cta gta aag aat tta gaa gaa ggg gtg cct ccc ttt     2016
Val Arg Asn Glu Leu Val Lys Asn Leu Glu Glu Gly Val Pro Pro Phe
            660                 665                 670

cag ctc tgc ctt cac tac aga gac ttt att ccc ggt gtg gcc att gct     2064
Gln Leu Cys Leu His Tyr Arg Asp Phe Ile Pro Gly Val Ala Ile Ala
        675                 680                 685

gcc aac atc atc cat gaa ggt ttc cat aaa agc cga aag gtg att gtt     2112
Ala Asn Ile Ile His Glu Gly Phe His Lys Ser Arg Lys Val Ile Val
    690                 695                 700

gtg gta tct cag cac ttc atc cag agc cgc tgg tgt atc ttt gaa tat     2160
Val Val Ser Gln His Phe Ile Gln Ser Arg Trp Cys Ile Phe Glu Tyr
705                 710                 715                 720

gag att gct cag acc tgg cag ttt ctg agc agt cgt gct ggt atc atc     2208
Glu Ile Ala Gln Thr Trp Gln Phe Leu Ser Ser Arg Ala Gly Ile Ile
                725                 730                 735

ttc att gtc ctg cag aag gtg gag aag tcc ctg ctc agg cag cag gtg     2256
Phe Ile Val Leu Gln Lys Val Glu Lys Ser Leu Leu Arg Gln Gln Val
            740                 745                 750

gag ctg tac cgc ctt ctc agc agg aac act tac ctg gag tgg gag gac     2304
Glu Leu Tyr Arg Leu Leu Ser Arg Asn Thr Tyr Leu Glu Trp Glu Asp
        755                 760                 765

agt gtc ctg ggg agg cac atc ttc tgg aga cga ctc aga aaa gcc ctg     2352
Ser Val Leu Gly Arg His Ile Phe Trp Arg Arg Leu Arg Lys Ala Leu
    770                 775                 780

ctg gat ggt aga ccg tgg aat cca gaa gga aca gtg ggt gca gga tgc     2400
Leu Asp Gly Arg Pro Trp Asn Pro Glu Gly Thr Val Gly Ala Gly Cys
785                 790                 795                 800

gaa tag                                                            2406
Glu
```

<210> 18

<211> 801
<212> PRT
<213> Saimiri sciureus

<220>
<221> misc_feature
<222> (14)..(14)
<223> The 'Xaa' at location 14 stands for Leu, or Phe.

<220>
<221> misc_feature
<222> (462)..(462)
<223> The 'Xaa' at location 462 stands for Thr.

<400> 18

Val Val Pro Asn Val Thr Tyr Gln Cys Met Glu Leu Asn Xaa Tyr Lys
1 5 10 15

Ile Pro Asp Asn Ile Pro Phe Ser Thr Lys Asn Leu Asp Leu Ser Phe
20 25 30

Asn Pro Leu Arg His Leu Gly Ser His Ser Phe Phe Asn Phe Pro Glu
35 40 45

Leu Gln Val Leu Asp Leu Ser Arg Cys Asp Ile Gln Thr Ile Glu Asp
50 55 60

Gly Ala Tyr Gln Ser Leu Ser His Leu Ser Thr Leu Ile Leu Thr Gly
65 70 75 80

Asn Pro Ile Gln Asn Leu Ala Leu Gly Ala Phe Ser Gly Leu Ser Ser
85 90 95

Leu Gln Lys Leu Val Ala Val Glu Thr His Leu Leu Ser Leu Glu Asn
100 105 110

Phe Pro Ile Gly His Leu Lys Thr Leu Lys Asp Leu Asn Val Ala His
115 120 125

Asn Leu Ile Gln Ser Phe Lys Leu Pro Glu Tyr Phe Ser Asn Leu Thr
130 135 140

Asn Leu Glu His Leu Asp Leu Ser Ser Asn Asn Ile Gln Asn Ile Tyr
145 150 155 160

Cys Lys Asp Leu Gln Val Leu His Gln Met Pro Leu Leu Asn Leu Ser
165 170 175

```
Leu Asp Leu Ser Leu Asn Pro Ile Asn Phe Ile Gln Pro Gly Ala Phe
            180                 185                 190

Lys Glu Ile Arg Leu His Lys Leu Thr Leu Arg Asn Asn Phe Asp Ser
        195                 200                 205

Leu Asn Ala Met Lys Thr Cys Ile Gln Gly Leu Ala Gly Leu Glu Val
    210                 215                 220

His Arg Leu Val Leu Gly Glu Phe Arg Asn Glu Arg Asn Ile Glu Asp
225                 230                 235                 240

Phe Asp Lys Ser Ala Leu Glu Gly Leu Cys Asn Leu Thr Ile Asn Glu
                245                 250                 255

Phe Arg Leu Ala Tyr Leu Asp Asp Phe Leu Asp Asp Ile Ile Asp Leu
            260                 265                 270

Phe Asn Cys Leu Ala Asn Val Ser Ser Phe Ser Leu Val Asn Val His
        275                 280                 285

Ile Lys Arg Val Glu Asp Phe Ser Tyr Asn Phe Arg Trp Gln His Leu
    290                 295                 300

Glu Leu Val Asn Cys Val Phe Gln Gln Phe Pro Pro Leu Lys Leu Lys
305                 310                 315                 320

Ser Leu Lys Arg Leu Thr Phe Thr Ala Asn Lys Gly Arg Asn His Phe
                325                 330                 335

Ser Glu Val Asp Leu Pro Ser Leu Glu Phe Leu Asp Leu Ser Arg Asn
            340                 345                 350

Gly Leu Ser Phe Lys Gly Cys Cys Ser Gln Ser Asp Phe Gly Thr Thr
        355                 360                 365

Ser Leu Lys Tyr Leu Asp Leu Ser Phe Asn Asp Val Ile Thr Met Gly
    370                 375                 380

Ser Asn Phe Leu Gly Leu Glu Gln Leu Glu His Leu Asp Phe Gln His
385                 390                 395                 400

Ser Asn Leu Lys Gln Met Ser Glu Phe Ser Val Phe Leu Ser Leu Arg
                405                 410                 415
```

Asn Leu Ile Tyr Leu Asp Ile Ser His Thr His Thr Arg Val Ala Phe
420 425 430

Asn Gly Ile Phe Asn Gly Leu Phe Ser Leu Lys Val Leu Lys Met Ala
435 440 445

Gly Asn Ser Phe Gln Gln Asn Phe Leu Glu Asp Ile Phe Xaa Asp Leu
450 455 460

Asn Asn Leu Ile Phe Leu Asp Leu Ser Glu Cys Gln Leu Glu Gln Leu
465 470 475 480

Ser Pro Thr Ala Phe Asp Ser Leu Pro Arg Leu Arg Ile Leu Asn Met
485 490 495

Ser His Asn Asn Phe Phe Ala Leu Asp Thr Phe Pro Tyr Lys His Leu
500 505 510

Tyr Ser Leu Gln Val Leu Asp Tyr Ser Leu Asn His Ile Gly Thr Ser
515 520 525

Lys Asn Gln Glu Leu Gln His Phe Pro Ser Ser Leu Ala Phe Leu Asn
530 535 540

Leu Thr Gln Asn Asp Phe Ala Cys Thr Cys Glu His Gln Ser Phe Leu
545 550 555 560

Gln Trp Ile Lys Asp Gln Arg Arg Leu Leu Val Glu Val Glu Gln Met
565 570 575

Glu Cys Ala Thr Pro Leu Asn Arg Lys Gly Ile Pro Val Leu Ser Leu
580 585 590

Asn Ile Thr Cys Gln Met Ser Lys Thr Ile Ile Gly Val Ser Val Leu
595 600 605

Ser Val Leu Val Val Ser Val Val Ala Val Leu Val Tyr Lys Phe Tyr
610 615 620

Phe His Leu Met Leu Leu Ala Gly Cys Ile Lys Tyr Gly Arg Gly Glu
625 630 635 640

Asn Thr Tyr Asp Ala Phe Val Ile Tyr Ser Ser Gln Asp Glu Asp Trp
645 650 655

Val Arg Asn Glu Leu Val Lys Asn Leu Glu Glu Gly Val Pro Pro Phe
660 665 670

Gln Leu Cys Leu His Tyr Arg Asp Phe Ile Pro Gly Val Ala Ile Ala
675 680 685

Ala Asn Ile Ile His Glu Gly Phe His Lys Ser Arg Lys Val Ile Val
690 695 700

Val Val Ser Gln His Phe Ile Gln Ser Arg Trp Cys Ile Phe Glu Tyr
705 710 715 720

Glu Ile Ala Gln Thr Trp Gln Phe Leu Ser Ser Arg Ala Gly Ile Ile
725 730 735

Phe Ile Val Leu Gln Lys Val Glu Lys Ser Leu Leu Arg Gln Gln Val
740 745 750

Glu Leu Tyr Arg Leu Leu Ser Arg Asn Thr Tyr Leu Glu Trp Glu Asp
755 760 765

Ser Val Leu Gly Arg His Ile Phe Trp Arg Arg Leu Arg Lys Ala Leu
770 775 780

Leu Asp Gly Arg Pro Trp Asn Pro Glu Gly Thr Val Gly Ala Gly Cys
785 790 795 800

Glu

<210> 19
<211> 2388
<212> DNA
<213> Papio hamadryas

<400> 19
gtggttccta acattactta tcaatgcatg gagctgaatt tctacaaaat ccccgacaac 60

atccccttct caaccaagaa cctggacctg agctttaatc ccctgaggca tttaggcagc 120

tatagcttcc tccgttttcc agaactgcag gtgctggatt tatccaggtg tgaaatccag 180

acaattgaag atggggcata tcagagccta agccacctct ccaccttaat attgacagga 240

aaccccatcc agagtttagc cctgggagcc ttttctggac tatcaagttt acagaagctg 300

gtggctgtgg agacaaatct agcatctcta gagaacttcc ccattggaca tctcaaaact 360

ttgaaagaac ttaatgtggc tcacaatctt atccagtctt tcaaattacc tgagtatttt 420

tctaatctga ccaatctaga gcacttggac ctttccagta acaagattca aaatatttat 480

tgcaaagact tgcaggttct acatcaaatg cccctaccca atctctcttt agacctgtcc 540

ctgaacccta taaactttat ccaaccaggt gcatttaaag aaattaggct tcataagctg 600

actttgagaa gtaatttga tgatttaaat gtaatgaaaa cttgtattca aggtctggct 660

ggtttagaag tccatcgttt ggttctggga gaatttagaa atgaaagaaa cttggaagag 720

tttgacaaat ctgctctgga gggattgtgc aatttgacca ttgaagaatt ccgattaaca 780

tacttagact actacctcga taatattatt gacttattta attgtttggc aaatgcttct 840

tcattttccc tggtgagtgt gaatattaaa agggtagaag acttttctta taatttcaga 900

tggcaacatt tagaattagt taactgtaaa tttgaacagt ttcccacatt ggaactcgaa 960

tctctcaaaa ggcttacttt cactgccaac aaaggtggga atgccttttc agaagttgat 1020

ctaccaagcc ttgagtttct agatctcagt agaaatggct tgagtttcaa aggttgctgt 1080

tctcaaagtg attttgggac aaccagccta aagtatttag atctgagctt caatgatgtt 1140

attaccatgg gttcaaactt cttgggctta gaacaactag aacatctgga tttccagcat 1200

tccaatttga aacagatgag tcaattttca gtattcctat cactcagaaa cctcatttac 1260

cttgacattt ctcatactca caccacagtt gctttcaatg gcattttcga tggcttgctc 1320

agtctcaaag tcttaaaaat ggctggcaat tctttccagg aaaacttcct tccagatatc 1380

ttcacagatc tgaaaaactt gaccttcctg gacctctctc agtgtcaact ggagcagttg 1440

tctccaacag catttgacac actcaacaag cttcaggtac taaatatgag ccacaacaac 1500

ttcttttcat tggatgtgtt tccttataag tgtctgccct ccctccaggt tctcgattac 1560

agtctcaatc acataatgac ttccaaaaac caggaacctc agcattttcc aagtagtcta 1620

gctttcttaa atcttactca gaatgacttt gcttgtactt gtgaacacca gagtttcctg 1680

cagtggatca aggaccagag gcagctcttg gtggaagctg aacgaatgga atgtgcaaca 1740

ccttcagata aacagggcat gcctgtgctg agtgtgaata ttacctgtca gatgaataag 1800

accatcattg gtgtgtctgt gttcagtgtg cttgtggtat ctgttgtagc agttctggtc 1860

tataagttct attttcacct gatgcttctt gctggctgca taaagtatgg tagaggtgaa 1920

aacatctatg atgcctttgt tatctactca agccaggatg aggactgggt aaggaatgag 1980

ctagtaaaga atttagaaga aggggtgcct ccctttcagc tctgccttca ctacagagac 2040

tttattcccg gtgtggccat tgctgcaaac atcatccatg aaggtttcca taaaagccga 2100

aaggtgattg ttgtggtgtc ccagcacttc atccagagcc gctggtgtat ctttgaatat 2160

gagattgctc agacctggca gtttctgagc agtcgtgcag gcataatctt cattgtcctg 2220

cagaaggtgg agaagaccct gctcaggcag caggtggagc tgtaccgcct tctcagcagg 2280

aacacttacc tggagtggga ggacagtgtc ctagggcagc acatcttctg gagacgactc 2340

agaaaagccc tgttggatgg cagatcgtgg aatccagaag aacagtag 2388

<210> 20
<211> 2388
<212> DNA
<213> Papio hamadryas

<220>
<221> CDS
<222> (1)..(2388)

<400> 20

```
gtg gtt cct aac att act tat caa tgc atg gag ctg aat ttc tac aaa      48
Val Val Pro Asn Ile Thr Tyr Gln Cys Met Glu Leu Asn Phe Tyr Lys
1               5                  10                  15

atc ccc gac aac atc ccc ttc tca acc aag aac ctg gac ctg agc ttt      96
Ile Pro Asp Asn Ile Pro Phe Ser Thr Lys Asn Leu Asp Leu Ser Phe
            20                  25                  30

aat ccc ctg agg cat tta ggc agc tat agc ttc ctc cgt ttt cca gaa     144
Asn Pro Leu Arg His Leu Gly Ser Tyr Ser Phe Leu Arg Phe Pro Glu
        35                  40                  45

ctg cag gtg ctg gat tta tcc agg tgt gaa atc cag aca att gaa gat     192
Leu Gln Val Leu Asp Leu Ser Arg Cys Glu Ile Gln Thr Ile Glu Asp
    50                  55                  60

ggg gca tat cag agc cta agc cac ctc tcc acc tta ata ttg aca gga     240
Gly Ala Tyr Gln Ser Leu Ser His Leu Ser Thr Leu Ile Leu Thr Gly
65                  70                  75                  80

aac ccc atc cag agt tta gcc ctg gga gcc ttt tct gga cta tca agt     288
Asn Pro Ile Gln Ser Leu Ala Leu Gly Ala Phe Ser Gly Leu Ser Ser
                85                  90                  95

tta cag aag ctg gtg gct gtg gag aca aat cta gca tct cta gag aac     336
Leu Gln Lys Leu Val Ala Val Glu Thr Asn Leu Ala Ser Leu Glu Asn
            100                 105                 110

ttc ccc att gga cat ctc aaa act ttg aaa gaa ctt aat gtg gct cac     384
Phe Pro Ile Gly His Leu Lys Thr Leu Lys Glu Leu Asn Val Ala His
        115                 120                 125

aat ctt atc cag tct ttc aaa tta cct gag tat ttt tct aat ctg acc     432
Asn Leu Ile Gln Ser Phe Lys Leu Pro Glu Tyr Phe Ser Asn Leu Thr
    130                 135                 140

aat cta gag cac ttg gac ctt tcc agt aac aag att caa aat att tat     480
Asn Leu Glu His Leu Asp Leu Ser Ser Asn Lys Ile Gln Asn Ile Tyr
145                 150                 155                 160

tgc aaa gac ttg cag gtt cta cat caa atg ccc cta ccc aat ctc tct     528
Cys Lys Asp Leu Gln Val Leu His Gln Met Pro Leu Pro Asn Leu Ser
```

```
                165                 170                 175

tta gac ctg tcc ctg aac cct ata aac ttt atc caa cca ggt gca ttt      576
Leu Asp Leu Ser Leu Asn Pro Ile Asn Phe Ile Gln Pro Gly Ala Phe
            180                 185                 190

aaa gaa att agg ctt cat aag ctg act ttg aga agt aat ttt gat gat      624
Lys Glu Ile Arg Leu His Lys Leu Thr Leu Arg Ser Asn Phe Asp Asp
        195                 200                 205

tta aat gta atg aaa act tgt att caa ggt ctg gct ggt tta gaa gtc      672
Leu Asn Val Met Lys Thr Cys Ile Gln Gly Leu Ala Gly Leu Glu Val
    210                 215                 220

cat cgt ttg gtt ctg gga gaa ttt aga aat gaa aga aac ttg gaa gag      720
His Arg Leu Val Leu Gly Glu Phe Arg Asn Glu Arg Asn Leu Glu Glu
225                 230                 235                 240

ttt gac aaa tct gct ctg gag gga ttg tgc aat ttg acc att gaa gaa      768
Phe Asp Lys Ser Ala Leu Glu Gly Leu Cys Asn Leu Thr Ile Glu Glu
                245                 250                 255

ttc cga tta aca tac tta gac tac tac ctc gat aat att att gac tta      816
Phe Arg Leu Thr Tyr Leu Asp Tyr Tyr Leu Asp Asn Ile Ile Asp Leu
            260                 265                 270

ttt aat tgt ttg gca aat gct tct tca ttt tcc ctg gtg agt gtg aat      864
Phe Asn Cys Leu Ala Asn Ala Ser Ser Phe Ser Leu Val Ser Val Asn
        275                 280                 285

att aaa agg gta gaa gac ttt tct tat aat ttc aga tgg caa cat tta      912
Ile Lys Arg Val Glu Asp Phe Ser Tyr Asn Phe Arg Trp Gln His Leu
    290                 295                 300

gaa tta gtt aac tgt aaa ttt gaa cag ttt ccc aca ttg gaa ctc gaa      960
Glu Leu Val Asn Cys Lys Phe Glu Gln Phe Pro Thr Leu Glu Leu Glu
305                 310                 315                 320

tct ctc aaa agg ctt act ttc act gcc aac aaa ggt ggg aat gcc ttt     1008
Ser Leu Lys Arg Leu Thr Phe Thr Ala Asn Lys Gly Gly Asn Ala Phe
                325                 330                 335

tca gaa gtt gat cta cca agc ctt gag ttt cta gat ctc agt aga aat     1056
Ser Glu Val Asp Leu Pro Ser Leu Glu Phe Leu Asp Leu Ser Arg Asn
            340                 345                 350

ggc ttg agt ttc aaa ggt tgc tgt tct caa agt gat ttt ggg aca acc     1104
Gly Leu Ser Phe Lys Gly Cys Cys Ser Gln Ser Asp Phe Gly Thr Thr
        355                 360                 365

agc cta aag tat tta gat ctg agc ttc aat gat gtt att acc atg ggt     1152
Ser Leu Lys Tyr Leu Asp Leu Ser Phe Asn Asp Val Ile Thr Met Gly
    370                 375                 380

tca aac ttc ttg ggc tta gaa caa cta gaa cat ctg gat ttc cag cat     1200
Ser Asn Phe Leu Gly Leu Glu Gln Leu Glu His Leu Asp Phe Gln His
385                 390                 395                 400

tcc aat ttg aaa cag atg agt caa ttt tca gta ttc cta tca ctc aga     1248
Ser Asn Leu Lys Gln Met Ser Gln Phe Ser Val Phe Leu Ser Leu Arg
```

```
                405                 410                 415

aac ctc att tac ctt gac att tct cat act cac acc aca gtt gct ttc     1296
Asn Leu Ile Tyr Leu Asp Ile Ser His Thr His Thr Thr Val Ala Phe
            420                 425                 430

aat ggc att ttc gat ggc ttg ctc agt ctc aaa gtc tta aaa atg gct     1344
Asn Gly Ile Phe Asp Gly Leu Leu Ser Leu Lys Val Leu Lys Met Ala
        435                 440                 445

ggc aat tct ttc cag gaa aac ttc ctt cca gat atc ttc aca gat ctg     1392
Gly Asn Ser Phe Gln Glu Asn Phe Leu Pro Asp Ile Phe Thr Asp Leu
    450                 455                 460

aaa aac ttg acc ttc ctg gac ctc tct cag tgt caa ctg gag cag ttg     1440
Lys Asn Leu Thr Phe Leu Asp Leu Ser Gln Cys Gln Leu Glu Gln Leu
465                 470                 475                 480

tct cca aca gca ttt gac aca ctc aac aag ctt cag gta cta aat atg     1488
Ser Pro Thr Ala Phe Asp Thr Leu Asn Lys Leu Gln Val Leu Asn Met
                485                 490                 495

agc cac aac aac ttc ttt tca ttg gat gtg ttt cct tat aag tgt ctg     1536
Ser His Asn Asn Phe Phe Ser Leu Asp Val Phe Pro Tyr Lys Cys Leu
            500                 505                 510

ccc tcc ctc cag gtt ctc gat tac agt ctc aat cac ata atg act tcc     1584
Pro Ser Leu Gln Val Leu Asp Tyr Ser Leu Asn His Ile Met Thr Ser
        515                 520                 525

aaa aac cag gaa cct cag cat ttt cca agt agt cta gct ttc tta aat     1632
Lys Asn Gln Glu Pro Gln His Phe Pro Ser Ser Leu Ala Phe Leu Asn
    530                 535                 540

ctt act cag aat gac ttt gct tgt act tgt gaa cac cag agt ttc ctg     1680
Leu Thr Gln Asn Asp Phe Ala Cys Thr Cys Glu His Gln Ser Phe Leu
545                 550                 555                 560

cag tgg atc aag gac cag agg cag ctc ttg gtg gaa gct gaa cga atg     1728
Gln Trp Ile Lys Asp Gln Arg Gln Leu Leu Val Glu Ala Glu Arg Met
                565                 570                 575

gaa tgt gca aca cct tca gat aaa cag ggc atg cct gtg ctg agt gtg     1776
Glu Cys Ala Thr Pro Ser Asp Lys Gln Gly Met Pro Val Leu Ser Val
            580                 585                 590

aat att acc tgt cag atg aat aag acc atc att ggt gtg tct gtg ttc     1824
Asn Ile Thr Cys Gln Met Asn Lys Thr Ile Ile Gly Val Ser Val Phe
        595                 600                 605

agt gtg ctt gtg gta tct gtt gta gca gtt ctg gtc tat aag ttc tat     1872
Ser Val Leu Val Val Ser Val Val Ala Val Leu Val Tyr Lys Phe Tyr
    610                 615                 620

ttt cac ctg atg ctt ctt gct ggc tgc ata aag tat ggt aga ggt gaa     1920
Phe His Leu Met Leu Leu Ala Gly Cys Ile Lys Tyr Gly Arg Gly Glu
625                 630                 635                 640

aac atc tat gat gcc ttt gtt atc tac tca agc cag gat gag gac tgg     1968
Asn Ile Tyr Asp Ala Phe Val Ile Tyr Ser Ser Gln Asp Glu Asp Trp
```

```
                645                 650                 655

gta agg aat gag cta gta aag aat tta gaa gaa ggg gtg cct ccc ttt      2016
Val Arg Asn Glu Leu Val Lys Asn Leu Glu Glu Gly Val Pro Pro Phe
            660                 665                 670

cag ctc tgc ctt cac tac aga gac ttt att ccc ggt gtg gcc att gct      2064
Gln Leu Cys Leu His Tyr Arg Asp Phe Ile Pro Gly Val Ala Ile Ala
        675                 680                 685

gca aac atc atc cat gaa ggt ttc cat aaa agc cga aag gtg att gtt      2112
Ala Asn Ile Ile His Glu Gly Phe His Lys Ser Arg Lys Val Ile Val
    690                 695                 700

gtg gtg tcc cag cac ttc atc cag agc cgc tgg tgt atc ttt gaa tat      2160
Val Val Ser Gln His Phe Ile Gln Ser Arg Trp Cys Ile Phe Glu Tyr
705                 710                 715                 720

gag att gct cag acc tgg cag ttt ctg agc agt cgt gca ggc ata atc      2208
Glu Ile Ala Gln Thr Trp Gln Phe Leu Ser Ser Arg Ala Gly Ile Ile
                725                 730                 735

ttc att gtc ctg cag aag gtg gag aag acc ctg ctc agg cag cag gtg      2256
Phe Ile Val Leu Gln Lys Val Glu Lys Thr Leu Leu Arg Gln Gln Val
            740                 745                 750

gag ctg tac cgc ctt ctc agc agg aac act tac ctg gag tgg gag gac      2304
Glu Leu Tyr Arg Leu Leu Ser Arg Asn Thr Tyr Leu Glu Trp Glu Asp
        755                 760                 765

agt gtc cta ggg cag cac atc ttc tgg aga cga ctc aga aaa gcc ctg      2352
Ser Val Leu Gly Gln His Ile Phe Trp Arg Arg Leu Arg Lys Ala Leu
    770                 775                 780

ttg gat ggc aga tcg tgg aat cca gaa gaa cag tag                      2388
Leu Asp Gly Arg Ser Trp Asn Pro Glu Glu Gln
785                 790                 795


<210>  21
<211>  795
<212>  PRT
<213>  Papio hamadryas

<400>  21

Val Val Pro Asn Ile Thr Tyr Gln Cys Met Glu Leu Asn Phe Tyr Lys
1               5                   10                  15


Ile Pro Asp Asn Ile Pro Phe Ser Thr Lys Asn Leu Asp Leu Ser Phe
            20                  25                  30


Asn Pro Leu Arg His Leu Gly Ser Tyr Ser Phe Leu Arg Phe Pro Glu
        35                  40                  45


Leu Gln Val Leu Asp Leu Ser Arg Cys Glu Ile Gln Thr Ile Glu Asp
    50                  55                  60
```

Gly Ala Tyr Gln Ser Leu Ser His Leu Ser Thr Leu Ile Leu Thr Gly
65 70 75 80

Asn Pro Ile Gln Ser Leu Ala Leu Gly Ala Phe Ser Gly Leu Ser Ser
85 90 95

Leu Gln Lys Leu Val Ala Val Glu Thr Asn Leu Ala Ser Leu Glu Asn
100 105 110

Phe Pro Ile Gly His Leu Lys Thr Leu Lys Glu Leu Asn Val Ala His
115 120 125

Asn Leu Ile Gln Ser Phe Lys Leu Pro Glu Tyr Phe Ser Asn Leu Thr
130 135 140

Asn Leu Glu His Leu Asp Leu Ser Ser Asn Lys Ile Gln Asn Ile Tyr
145 150 155 160

Cys Lys Asp Leu Gln Val Leu His Gln Met Pro Leu Pro Asn Leu Ser
165 170 175

Leu Asp Leu Ser Leu Asn Pro Ile Asn Phe Ile Gln Pro Gly Ala Phe
180 185 190

Lys Glu Ile Arg Leu His Lys Leu Thr Leu Arg Ser Asn Phe Asp Asp
195 200 205

Leu Asn Val Met Lys Thr Cys Ile Gln Gly Leu Ala Gly Leu Glu Val
210 215 220

His Arg Leu Val Leu Gly Glu Phe Arg Asn Glu Arg Asn Leu Glu Glu
225 230 235 240

Phe Asp Lys Ser Ala Leu Glu Gly Leu Cys Asn Leu Thr Ile Glu Glu
245 250 255

Phe Arg Leu Thr Tyr Leu Asp Tyr Tyr Leu Asp Asn Ile Ile Asp Leu
260 265 270

Phe Asn Cys Leu Ala Asn Ala Ser Ser Phe Ser Leu Val Ser Val Asn
275 280 285

Ile Lys Arg Val Glu Asp Phe Ser Tyr Asn Phe Arg Trp Gln His Leu
290 295 300

Glu Leu Val Asn Cys Lys Phe Glu Gln Phe Pro Thr Leu Glu Leu Glu
305 310 315 320

Ser Leu Lys Arg Leu Thr Phe Thr Ala Asn Lys Gly Gly Asn Ala Phe
325 330 335

Ser Glu Val Asp Leu Pro Ser Leu Glu Phe Leu Asp Leu Ser Arg Asn
340 345 350

Gly Leu Ser Phe Lys Gly Cys Cys Ser Gln Ser Asp Phe Gly Thr Thr
355 360 365

Ser Leu Lys Tyr Leu Asp Leu Ser Phe Asn Asp Val Ile Thr Met Gly
370 375 380

Ser Asn Phe Leu Gly Leu Glu Gln Leu Glu His Leu Asp Phe Gln His
385 390 395 400

Ser Asn Leu Lys Gln Met Ser Gln Phe Ser Val Phe Leu Ser Leu Arg
405 410 415

Asn Leu Ile Tyr Leu Asp Ile Ser His Thr His Thr Thr Val Ala Phe
420 425 430

Asn Gly Ile Phe Asp Gly Leu Leu Ser Leu Lys Val Leu Lys Met Ala
435 440 445

Gly Asn Ser Phe Gln Glu Asn Phe Leu Pro Asp Ile Phe Thr Asp Leu
450 455 460

Lys Asn Leu Thr Phe Leu Asp Leu Ser Gln Cys Gln Leu Glu Gln Leu
465 470 475 480

Ser Pro Thr Ala Phe Asp Thr Leu Asn Lys Leu Gln Val Leu Asn Met
485 490 495

Ser His Asn Asn Phe Phe Ser Leu Asp Val Phe Pro Tyr Lys Cys Leu
500 505 510

Pro Ser Leu Gln Val Leu Asp Tyr Ser Leu Asn His Ile Met Thr Ser
515 520 525

Lys Asn Gln Glu Pro Gln His Phe Pro Ser Ser Leu Ala Phe Leu Asn
530 535 540

```
Leu Thr Gln Asn Asp Phe Ala Cys Thr Cys Glu His Gln Ser Phe Leu
545                 550                 555                 560

Gln Trp Ile Lys Asp Gln Arg Gln Leu Leu Val Glu Ala Glu Arg Met
                565                 570                 575

Glu Cys Ala Thr Pro Ser Asp Lys Gln Gly Met Pro Val Leu Ser Val
            580                 585                 590

Asn Ile Thr Cys Gln Met Asn Lys Thr Ile Ile Gly Val Ser Val Phe
        595                 600                 605

Ser Val Leu Val Val Ser Val Val Ala Val Leu Val Tyr Lys Phe Tyr
    610                 615                 620

Phe His Leu Met Leu Leu Ala Gly Cys Ile Lys Tyr Gly Arg Gly Glu
625                 630                 635                 640

Asn Ile Tyr Asp Ala Phe Val Ile Tyr Ser Ser Gln Asp Glu Asp Trp
                645                 650                 655

Val Arg Asn Glu Leu Val Lys Asn Leu Glu Glu Gly Val Pro Pro Phe
            660                 665                 670

Gln Leu Cys Leu His Tyr Arg Asp Phe Ile Pro Gly Val Ala Ile Ala
        675                 680                 685

Ala Asn Ile Ile His Glu Gly Phe His Lys Ser Arg Lys Val Ile Val
    690                 695                 700

Val Val Ser Gln His Phe Ile Gln Ser Arg Trp Cys Ile Phe Glu Tyr
705                 710                 715                 720

Glu Ile Ala Gln Thr Trp Gln Phe Leu Ser Ser Arg Ala Gly Ile Ile
                725                 730                 735

Phe Ile Val Leu Gln Lys Val Glu Lys Thr Leu Leu Arg Gln Gln Val
            740                 745                 750

Glu Leu Tyr Arg Leu Leu Ser Arg Asn Thr Tyr Leu Glu Trp Glu Asp
        755                 760                 765

Ser Val Leu Gly Gln His Ile Phe Trp Arg Arg Leu Arg Lys Ala Leu
    770                 775                 780
```

Leu Asp Gly Arg Ser Trp Asn Pro Glu Glu Gln
785 790 795

<210> 22
<211> 2427
<212> DNA
<213> Pan troglodytes

<400> 22
gtggttccta atattactta tcaatgcatg gagctgaatt tctacaaaat ccccgacaac 60

ctccccttct caaccaagaa cctggacctg agctttaatc ccctgaggca tttaggcagc 120

tatagcttct tcagtttccc agaactgcag gtgctggatt tatccaggtg tgaaatccag 180

acaattgaag atggggcata tcagagccta agccacctct ccaccttaat attgacagga 240

aaccccatcc agagtttagc cctgggagcc ttttctggac tatcaagttt acagaagctg 300

gtggctgtgg agacaaatct agcatctcta gagaacttcc ccattggaca tctcaaaact 360

ttgaaagaac ttaatgtggc tcacaatctt atccaatctt tcaaattacc tgagtatttt 420

tctaatctga ccaatctaga gcacttggac ctttccagca acaagattca aagtatttat 480

tgcacagact tgcgggttct acatcaaatg cccctactca atctctcttt agacctgtcc 540

ctgaacccta tgaactttat ccaaccaggt gcatttaaag aaattaggct tcataagctg 600

actttgagaa ataattttga tagtttaaat gtaatgaaaa cttgtattca aggtctggct 660

ggtttagaag tccatcgttt ggttctggga gaatttagaa atgaagaaaa cttggaaaag 720

tttgacaaat ctgctctaga gggcctgtgc aatttgacca ttgaagaatt ccgattagca 780

tacttagact actacctcga tgatattatt gacttattta attgtttgac aaatgtttct 840

tcattttccc tggtgagtgt gactattaaa agcgtaaaag acttttctta taatttcgga 900

tggcaacatt tagaattagt taagtgtaaa tttggacagt ttcccacatt gaaactcaaa 960

tctctcaaaa ggcttacttt cacttccaac aaaggtggga atgctttttc agaagttgat 1020

ctaccaagcc ttgagtttct agatctcagt agaaatggct tgagtttcaa aggttgctgt 1080

tctcaaagtg attttgggac aaccagccta aagtatttag atctgagctt caatggtgtt 1140

attaccatga gttcaaactt cttgggctta gaacaactag aacatctgga tttccagcat 1200

tccaatttga aacaaatgag tgagttttca gtattcctat cactcagaaa cctcatttac 1260

cttgacattt ctcatactca caccagagtt gctttcaatg gcatcttcaa tggcttgtcc 1320

agtctcgaag tcttgaaaat ggctggcaat tctttccagg aaaacttcct tccagatatc 1380

ttcacagagc tgagaaactt gaccttcctg gacctctctc agtgtcaact ggagcagttg 1440

tctccaacag catttaactc actctccagt cttcaggtac taaatatgag ccacaacaac 1500

ttcttttcat tggatacgtt tccttataag tgtctgaact ccctccaggt tcttgattac 1560

agtctcaatc acataatgac ttccaaaaaa caggaactac agcattttcc aagtagtcta 1620

gctttcttaa atcttactca gaatgacttt gcttgtactt gtgaacacca aagtttcctg 1680

caatggatca aggaccagag gcagctcttg gtggaagttg aacgaatgga atgtgcaaca 1740

ccttcagata agcagggcat gcctgtgctg agtttgaata tcacctgtca gatgaataag 1800

accatcattg gtgtgtcggt cctcagtgtg cttgtagtat ctgttgtagc agttctggtc 1860

tataagttct attttcacct gatgcttctt gctggctgca taaagtatgg tagaggtgaa 1920

aacatctatg atgcctttgt tatctactca agccaggatg aggactgggt aaggaatgag 1980

ctagtaaaga atttagaaga aggggtgcct ccatttcagc tctgccttca ctacagagac 2040

tttattcccg gtgtggccat tgctgccaac atcatccatg aaggtttcca taaaagccga 2100

aaggtgattg ttgtggtgtc ccagcacttc atccagagcc gctggtgtat ctttgaatat 2160

gagattgctc agacgtggca gtttctgagc agtcgtgctg gtatcatctt cattgtcctg 2220

cagaaggtgg agaagaccct gctcaggcgg caggtggagc tgtaccgcct tctyagcagg 2280

aacacttacc tggagtggga ggacagtgtc ctggggcggc acatcttctg gagacgactc 2340

agaaaagccc tgctggatgg taaatcatgg aatccagaag gaacagtggg tacaggatgc 2400

aattggcagg aagcaacatc tatctga 2427

<210> 23
<211> 2427
<212> DNA
<213> Pan troglodytes

<220>
<221> CDS
<222> (1)..(2427)

<400> 23

```
gtg gtt cct aat att act tat caa tgc atg gag ctg aat ttc tac aaa      48
Val Val Pro Asn Ile Thr Tyr Gln Cys Met Glu Leu Asn Phe Tyr Lys
1               5                  10                  15

atc ccc gac aac ctc ccc ttc tca acc aag aac ctg gac ctg agc ttt      96
Ile Pro Asp Asn Leu Pro Phe Ser Thr Lys Asn Leu Asp Leu Ser Phe
            20                  25                  30

aat ccc ctg agg cat tta ggc agc tat agc ttc ttc agt ttc cca gaa     144
Asn Pro Leu Arg His Leu Gly Ser Tyr Ser Phe Phe Ser Phe Pro Glu
        35                  40                  45

ctg cag gtg ctg gat tta tcc agg tgt gaa atc cag aca att gaa gat     192
Leu Gln Val Leu Asp Leu Ser Arg Cys Glu Ile Gln Thr Ile Glu Asp
```

```
    50                  55                  60

ggg gca tat cag agc cta agc cac ctc tcc acc tta ata ttg aca gga      240
Gly Ala Tyr Gln Ser Leu Ser His Leu Ser Thr Leu Ile Leu Thr Gly
65                  70                  75                  80

aac ccc atc cag agt tta gcc ctg gga gcc ttt tct gga cta tca agt      288
Asn Pro Ile Gln Ser Leu Ala Leu Gly Ala Phe Ser Gly Leu Ser Ser
                85                  90                  95

tta cag aag ctg gtg gct gtg gag aca aat cta gca tct cta gag aac      336
Leu Gln Lys Leu Val Ala Val Glu Thr Asn Leu Ala Ser Leu Glu Asn
            100                 105                 110

ttc ccc att gga cat ctc aaa act ttg aaa gaa ctt aat gtg gct cac      384
Phe Pro Ile Gly His Leu Lys Thr Leu Lys Glu Leu Asn Val Ala His
        115                 120                 125

aat ctt atc caa tct ttc aaa tta cct gag tat ttt tct aat ctg acc      432
Asn Leu Ile Gln Ser Phe Lys Leu Pro Glu Tyr Phe Ser Asn Leu Thr
    130                 135                 140

aat cta gag cac ttg gac ctt tcc agc aac aag att caa agt att tat      480
Asn Leu Glu His Leu Asp Leu Ser Ser Asn Lys Ile Gln Ser Ile Tyr
145                 150                 155                 160

tgc aca gac ttg cgg gtt cta cat caa atg ccc cta ctc aat ctc tct      528
Cys Thr Asp Leu Arg Val Leu His Gln Met Pro Leu Leu Asn Leu Ser
                165                 170                 175

tta gac ctg tcc ctg aac cct atg aac ttt atc caa cca ggt gca ttt      576
Leu Asp Leu Ser Leu Asn Pro Met Asn Phe Ile Gln Pro Gly Ala Phe
            180                 185                 190

aaa gaa att agg ctt cat aag ctg act ttg aga aat aat ttt gat agt      624
Lys Glu Ile Arg Leu His Lys Leu Thr Leu Arg Asn Asn Phe Asp Ser
        195                 200                 205

tta aat gta atg aaa act tgt att caa ggt ctg gct ggt tta gaa gtc      672
Leu Asn Val Met Lys Thr Cys Ile Gln Gly Leu Ala Gly Leu Glu Val
    210                 215                 220

cat cgt ttg gtt ctg gga gaa ttt aga aat gaa gaa aac ttg gaa aag      720
His Arg Leu Val Leu Gly Glu Phe Arg Asn Glu Glu Asn Leu Glu Lys
225                 230                 235                 240

ttt gac aaa tct gct cta gag ggc ctg tgc aat ttg acc att gaa gaa      768
Phe Asp Lys Ser Ala Leu Glu Gly Leu Cys Asn Leu Thr Ile Glu Glu
                245                 250                 255

ttc cga tta gca tac tta gac tac tac ctc gat gat att att gac tta      816
Phe Arg Leu Ala Tyr Leu Asp Tyr Tyr Leu Asp Asp Ile Ile Asp Leu
            260                 265                 270

ttt aat tgt ttg aca aat gtt tct tca ttt tcc ctg gtg agt gtg act      864
Phe Asn Cys Leu Thr Asn Val Ser Ser Phe Ser Leu Val Ser Val Thr
        275                 280                 285

att aaa agc gta aaa gac ttt tct tat aat ttc gga tgg caa cat tta      912
Ile Lys Ser Val Lys Asp Phe Ser Tyr Asn Phe Gly Trp Gln His Leu
```

```
    290                 295                 300

gaa tta gtt aag tgt aaa ttt gga cag ttt ccc aca ttg aaa ctc aaa      960
Glu Leu Val Lys Cys Lys Phe Gly Gln Phe Pro Thr Leu Lys Leu Lys
305                 310                 315                 320

tct ctc aaa agg ctt act ttc act tcc aac aaa ggt ggg aat gct ttt     1008
Ser Leu Lys Arg Leu Thr Phe Thr Ser Asn Lys Gly Gly Asn Ala Phe
                325                 330                 335

tca gaa gtt gat cta cca agc ctt gag ttt cta gat ctc agt aga aat     1056
Ser Glu Val Asp Leu Pro Ser Leu Glu Phe Leu Asp Leu Ser Arg Asn
            340                 345                 350

ggc ttg agt ttc aaa ggt tgc tgt tct caa agt gat ttt ggg aca acc     1104
Gly Leu Ser Phe Lys Gly Cys Cys Ser Gln Ser Asp Phe Gly Thr Thr
        355                 360                 365

agc cta aag tat tta gat ctg agc ttc aat ggt gtt att acc atg agt     1152
Ser Leu Lys Tyr Leu Asp Leu Ser Phe Asn Gly Val Ile Thr Met Ser
    370                 375                 380

tca aac ttc ttg ggc tta gaa caa cta gaa cat ctg gat ttc cag cat     1200
Ser Asn Phe Leu Gly Leu Glu Gln Leu Glu His Leu Asp Phe Gln His
385                 390                 395                 400

tcc aat ttg aaa caa atg agt gag ttt tca gta ttc cta tca ctc aga     1248
Ser Asn Leu Lys Gln Met Ser Glu Phe Ser Val Phe Leu Ser Leu Arg
                405                 410                 415

aac ctc att tac ctt gac att tct cat act cac acc aga gtt gct ttc     1296
Asn Leu Ile Tyr Leu Asp Ile Ser His Thr His Thr Arg Val Ala Phe
            420                 425                 430

aat ggc atc ttc aat ggc ttg tcc agt ctc gaa gtc ttg aaa atg gct     1344
Asn Gly Ile Phe Asn Gly Leu Ser Ser Leu Glu Val Leu Lys Met Ala
        435                 440                 445

ggc aat tct ttc cag gaa aac ttc ctt cca gat atc ttc aca gag ctg     1392
Gly Asn Ser Phe Gln Glu Asn Phe Leu Pro Asp Ile Phe Thr Glu Leu
    450                 455                 460

aga aac ttg acc ttc ctg gac ctc tct cag tgt caa ctg gag cag ttg     1440
Arg Asn Leu Thr Phe Leu Asp Leu Ser Gln Cys Gln Leu Glu Gln Leu
465                 470                 475                 480

tct cca aca gca ttt aac tca ctc tcc agt ctt cag gta cta aat atg     1488
Ser Pro Thr Ala Phe Asn Ser Leu Ser Ser Leu Gln Val Leu Asn Met
                485                 490                 495

agc cac aac aac ttc ttt tca ttg gat acg ttt cct tat aag tgt ctg     1536
Ser His Asn Asn Phe Phe Ser Leu Asp Thr Phe Pro Tyr Lys Cys Leu
            500                 505                 510

aac tcc ctc cag gtt ctt gat tac agt ctc aat cac ata atg act tcc     1584
Asn Ser Leu Gln Val Leu Asp Tyr Ser Leu Asn His Ile Met Thr Ser
        515                 520                 525

aaa aaa cag gaa cta cag cat ttt cca agt agt cta gct ttc tta aat     1632
Lys Lys Gln Glu Leu Gln His Phe Pro Ser Ser Leu Ala Phe Leu Asn
```

```
    530                 535                 540

ctt act cag aat gac ttt gct tgt act tgt gaa cac caa agt ttc ctg     1680
Leu Thr Gln Asn Asp Phe Ala Cys Thr Cys Glu His Gln Ser Phe Leu
545                 550                 555                 560

caa tgg atc aag gac cag agg cag ctc ttg gtg gaa gtt gaa cga atg     1728
Gln Trp Ile Lys Asp Gln Arg Gln Leu Leu Val Glu Val Glu Arg Met
                565                 570                 575

gaa tgt gca aca cct tca gat aag cag ggc atg cct gtg ctg agt ttg     1776
Glu Cys Ala Thr Pro Ser Asp Lys Gln Gly Met Pro Val Leu Ser Leu
            580                 585                 590

aat atc acc tgt cag atg aat aag acc atc att ggt gtg tcg gtc ctc     1824
Asn Ile Thr Cys Gln Met Asn Lys Thr Ile Ile Gly Val Ser Val Leu
        595                 600                 605

agt gtg ctt gta gta tct gtt gta gca gtt ctg gtc tat aag ttc tat     1872
Ser Val Leu Val Val Ser Val Val Ala Val Leu Val Tyr Lys Phe Tyr
    610                 615                 620

ttt cac ctg atg ctt ctt gct ggc tgc ata aag tat ggt aga ggt gaa     1920
Phe His Leu Met Leu Leu Ala Gly Cys Ile Lys Tyr Gly Arg Gly Glu
625                 630                 635                 640

aac atc tat gat gcc ttt gtt atc tac tca agc cag gat gag gac tgg     1968
Asn Ile Tyr Asp Ala Phe Val Ile Tyr Ser Ser Gln Asp Glu Asp Trp
                645                 650                 655

gta agg aat gag cta gta aag aat tta gaa gaa ggg gtg cct cca ttt     2016
Val Arg Asn Glu Leu Val Lys Asn Leu Glu Glu Gly Val Pro Pro Phe
            660                 665                 670

cag ctc tgc ctt cac tac aga gac ttt att ccc ggt gtg gcc att gct     2064
Gln Leu Cys Leu His Tyr Arg Asp Phe Ile Pro Gly Val Ala Ile Ala
        675                 680                 685

gcc aac atc atc cat gaa ggt ttc cat aaa agc cga aag gtg att gtt     2112
Ala Asn Ile Ile His Glu Gly Phe His Lys Ser Arg Lys Val Ile Val
    690                 695                 700

gtg gtg tcc cag cac ttc atc cag agc cgc tgg tgt atc ttt gaa tat     2160
Val Val Ser Gln His Phe Ile Gln Ser Arg Trp Cys Ile Phe Glu Tyr
705                 710                 715                 720

gag att gct cag acg tgg cag ttt ctg agc agt cgt gct ggt atc atc     2208
Glu Ile Ala Gln Thr Trp Gln Phe Leu Ser Ser Arg Ala Gly Ile Ile
                725                 730                 735

ttc att gtc ctg cag aag gtg gag aag acc ctg ctc agg cgg cag gtg     2256
Phe Ile Val Leu Gln Lys Val Glu Lys Thr Leu Leu Arg Arg Gln Val
            740                 745                 750

gag ctg tac cgc ctt cty agc agg aac act tac ctg gag tgg gag gac     2304
Glu Leu Tyr Arg Leu Xaa Ser Arg Asn Thr Tyr Leu Glu Trp Glu Asp
        755                 760                 765

agt gtc ctg ggg cgg cac atc ttc tgg aga cga ctc aga aaa gcc ctg     2352
Ser Val Leu Gly Arg His Ile Phe Trp Arg Arg Leu Arg Lys Ala Leu
```

770 775 780

ctg gat ggt aaa tca tgg aat cca gaa gga aca gtg ggt aca gga tgc 2400
Leu Asp Gly Lys Ser Trp Asn Pro Glu Gly Thr Val Gly Thr Gly Cys
785 790 795 800

aat tgg cag gaa gca aca tct atc tga 2427
Asn Trp Gln Glu Ala Thr Ser Ile
805

<210> 24
<211> 808
<212> PRT
<213> Pan troglodytes

<220>
<221> misc_feature
<222> (758)..(758)
<223> The 'Xaa' at location 758 stands for Leu.

<400> 24

Val Val Pro Asn Ile Thr Tyr Gln Cys Met Glu Leu Asn Phe Tyr Lys
1 5 10 15

Ile Pro Asp Asn Leu Pro Phe Ser Thr Lys Asn Leu Asp Leu Ser Phe
20 25 30

Asn Pro Leu Arg His Leu Gly Ser Tyr Ser Phe Phe Ser Phe Pro Glu
35 40 45

Leu Gln Val Leu Asp Leu Ser Arg Cys Glu Ile Gln Thr Ile Glu Asp
50 55 60

Gly Ala Tyr Gln Ser Leu Ser His Leu Ser Thr Leu Ile Leu Thr Gly
65 70 75 80

Asn Pro Ile Gln Ser Leu Ala Leu Gly Ala Phe Ser Gly Leu Ser Ser
85 90 95

Leu Gln Lys Leu Val Ala Val Glu Thr Asn Leu Ala Ser Leu Glu Asn
100 105 110

Phe Pro Ile Gly His Leu Lys Thr Leu Lys Glu Leu Asn Val Ala His
115 120 125

Asn Leu Ile Gln Ser Phe Lys Leu Pro Glu Tyr Phe Ser Asn Leu Thr
130 135 140

Asn Leu Glu His Leu Asp Leu Ser Ser Asn Lys Ile Gln Ser Ile Tyr

145 150 155 160

Cys Thr Asp Leu Arg Val Leu His Gln Met Pro Leu Leu Asn Leu Ser
165 170 175

Leu Asp Leu Ser Leu Asn Pro Met Asn Phe Ile Gln Pro Gly Ala Phe
180 185 190

Lys Glu Ile Arg Leu His Lys Leu Thr Leu Arg Asn Asn Phe Asp Ser
195 200 205

Leu Asn Val Met Lys Thr Cys Ile Gln Gly Leu Ala Gly Leu Glu Val
210 215 220

His Arg Leu Val Leu Gly Glu Phe Arg Asn Glu Glu Asn Leu Glu Lys
225 230 235 240

Phe Asp Lys Ser Ala Leu Glu Gly Leu Cys Asn Leu Thr Ile Glu Glu
245 250 255

Phe Arg Leu Ala Tyr Leu Asp Tyr Tyr Leu Asp Asp Ile Ile Asp Leu
260 265 270

Phe Asn Cys Leu Thr Asn Val Ser Ser Phe Ser Leu Val Ser Val Thr
275 280 285

Ile Lys Ser Val Lys Asp Phe Ser Tyr Asn Phe Gly Trp Gln His Leu
290 295 300

Glu Leu Val Lys Cys Lys Phe Gly Gln Phe Pro Thr Leu Lys Leu Lys
305 310 315 320

Ser Leu Lys Arg Leu Thr Phe Thr Ser Asn Lys Gly Gly Asn Ala Phe
325 330 335

Ser Glu Val Asp Leu Pro Ser Leu Glu Phe Leu Asp Leu Ser Arg Asn
340 345 350

Gly Leu Ser Phe Lys Gly Cys Cys Ser Gln Ser Asp Phe Gly Thr Thr
355 360 365

Ser Leu Lys Tyr Leu Asp Leu Ser Phe Asn Gly Val Ile Thr Met Ser
370 375 380

Ser Asn Phe Leu Gly Leu Glu Gln Leu Glu His Leu Asp Phe Gln His

385 390 395 400

Ser Asn Leu Lys Gln Met Ser Glu Phe Ser Val Phe Leu Ser Leu Arg
405 410 415

Asn Leu Ile Tyr Leu Asp Ile Ser His Thr His Thr Arg Val Ala Phe
420 425 430

Asn Gly Ile Phe Asn Gly Leu Ser Ser Leu Glu Val Leu Lys Met Ala
435 440 445

Gly Asn Ser Phe Gln Glu Asn Phe Leu Pro Asp Ile Phe Thr Glu Leu
450 455 460

Arg Asn Leu Thr Phe Leu Asp Leu Ser Gln Cys Gln Leu Glu Gln Leu
465 470 475 480

Ser Pro Thr Ala Phe Asn Ser Leu Ser Ser Leu Gln Val Leu Asn Met
485 490 495

Ser His Asn Asn Phe Phe Ser Leu Asp Thr Phe Pro Tyr Lys Cys Leu
500 505 510

Asn Ser Leu Gln Val Leu Asp Tyr Ser Leu Asn His Ile Met Thr Ser
515 520 525

Lys Lys Gln Glu Leu Gln His Phe Pro Ser Ser Leu Ala Phe Leu Asn
530 535 540

Leu Thr Gln Asn Asp Phe Ala Cys Thr Cys Glu His Gln Ser Phe Leu
545 550 555 560

Gln Trp Ile Lys Asp Gln Arg Gln Leu Leu Val Glu Val Glu Arg Met
565 570 575

Glu Cys Ala Thr Pro Ser Asp Lys Gln Gly Met Pro Val Leu Ser Leu
580 585 590

Asn Ile Thr Cys Gln Met Asn Lys Thr Ile Ile Gly Val Ser Val Leu
595 600 605

Ser Val Leu Val Val Ser Val Val Ala Val Leu Val Tyr Lys Phe Tyr
610 615 620

Phe His Leu Met Leu Leu Ala Gly Cys Ile Lys Tyr Gly Arg Gly Glu

625 630 635 640

Asn Ile Tyr Asp Ala Phe Val Ile Tyr Ser Ser Gln Asp Glu Asp Trp
645 650 655

Val Arg Asn Glu Leu Val Lys Asn Leu Glu Glu Gly Val Pro Pro Phe
660 665 670

Gln Leu Cys Leu His Tyr Arg Asp Phe Ile Pro Gly Val Ala Ile Ala
675 680 685

Ala Asn Ile Ile His Glu Gly Phe His Lys Ser Arg Lys Val Ile Val
690 695 700

Val Val Ser Gln His Phe Ile Gln Ser Arg Trp Cys Ile Phe Glu Tyr
705 710 715 720

Glu Ile Ala Gln Thr Trp Gln Phe Leu Ser Ser Arg Ala Gly Ile Ile
725 730 735

Phe Ile Val Leu Gln Lys Val Glu Lys Thr Leu Leu Arg Arg Gln Val
740 745 750

Glu Leu Tyr Arg Leu Xaa Ser Arg Asn Thr Tyr Leu Glu Trp Glu Asp
755 760 765

Ser Val Leu Gly Arg His Ile Phe Trp Arg Arg Leu Arg Lys Ala Leu
770 775 780

Leu Asp Gly Lys Ser Trp Asn Pro Glu Gly Thr Val Gly Thr Gly Cys
785 790 795 800

Asn Trp Gln Glu Ala Thr Ser Ile
805

<210> 25
<211> 35
<212> PRT
<213> Amino acid

<400> 25

Cys Asn Leu Thr Ile Glu Glu Phe Arg Leu Thr Tyr Leu Asp Tyr Tyr
1 5 10 15

Leu Asp Asp Ile Ile Asp Leu Phe Asn Cys Leu Ala Asn Ala Ser Ser
20 25 30

Phe Ser Leu
35

<210> 26
<211> 35
<212> PRT
<213> Amino acid

<400> 26

Cys Asn Leu Thr Ile Glu Glu Phe Arg Leu Thr Tyr Leu Asp Tyr Tyr
1 5 10 15

Leu Asp Gly Ile Ile Asp Leu Phe Asn Cys Leu Ala Asn Ala Ser Ser
20 25 30

Phe Ser Leu
35

<210> 27
<211> 35
<212> PRT
<213> Amino acid

<400> 27

Cys Asn Leu Thr Ile Glu Glu Phe Arg Leu Thr Tyr Leu Asp Tyr Tyr
1 5 10 15

Leu Asp Asp Ile Ile Asp Leu Phe Asn Cys Leu Ala Asn Ala Ser Ser
20 25 30

Phe Ser Leu
35

<210> 28
<211> 35
<212> PRT
<213> Amino acid

<400> 28

Cys Asn Leu Thr Ile Glu Glu Phe Arg Leu Thr Tyr Leu Asp Tyr Tyr
1 5 10 15

Leu Asp Asp Ile Ile Asp Leu Phe Asn Cys Leu Ala Asn Ala Ser Ser
20 25 30

Phe Ser Leu

35

<210> 29
<211> 35
<212> PRT
<213> Amino acid

<400> 29

Cys Asn Leu Thr Ile Glu Glu Phe Arg Leu Thr Tyr Leu Asp Tyr Tyr
1 5 10 15

Leu Asp Asp Ile Ile Asp Leu Phe Asn Cys Leu Ala Asn Ala Ser Ser
20 25 30

Phe Ser Leu
35

<210> 30
<211> 35
<212> PRT
<213> Amino acid

<400> 30

Cys Asn Leu Thr Ile Glu Glu Phe Arg Leu Ala Tyr Leu Asp Tyr Tyr
1 5 10 15

Leu Asp Asp Ile Ile Asp Leu Phe Asn Cys Leu Ala Asn Val Ser Ser
20 25 30

Phe Ser Leu
35

<210> 31
<211> 35
<212> PRT
<213> Amino acid

<400> 31

Cys Asn Leu Thr Ile Glu Glu Phe Arg Leu Thr Tyr Leu Asp Tyr Tyr
1 5 10 15

Leu Asp Asp Ile Ile Asp Leu Phe Asn Cys Leu Ala Asn Ala Ser Ser
20 25 30

Phe Ser Leu
35

<210> 32
<211> 35
<212> PRT
<213> Amino acid

<400> 32

Cys Asn Leu Thr Ile Glu Glu Phe Arg Leu Thr Tyr Leu Asp Tyr Tyr
1 5 10 15

Leu Asp Asn Ile Ile Asp Leu Phe Asn Cys Leu Ala Asn Ala Ser Ser
20 25 30

Phe Ser Leu
35

<210> 33
<211> 35
<212> PRT
<213> Amino acid

<400> 33

Cys Asn Leu Thr Ile Glu Glu Phe Arg Leu Thr Tyr Leu Asp Tyr Tyr
1 5 10 15

Leu Asp Asn Ile Ile Asp Leu Phe Asn Cys Leu Ala Asn Ala Ser Ser
20 25 30

Phe Ser Leu
35

<210> 34
<211> 36
<212> PRT
<213> Amino acid

<400> 34

His Asn Leu Thr Ile Glu Glu Phe Arg Leu Ala Tyr Ile Asp Asn Tyr
1 5 10 15

Ser Ser Lys Asp Ser Ile Asp Leu Leu Asn Cys Leu Ala Asp Ile Ser
20 25 30

Lys Ile Ser Leu
35

<210> 35
<211> 35
<212> PRT

<213> Amino acid

<400> 35

Cys Asn Leu Thr Ile Glu Gln Phe Arg Ile Ala Tyr Leu Asp Lys Phe
1 5 10 15

Ser Gly Asp Asp Thr Asp Leu Phe Asn Cys Leu Ala Asn Val Ser Val
20 25 30

Ile Ser Leu
35

<210> 36
<211> 35
<212> PRT
<213> Amino acid

<400> 36

Cys Asn Leu Ile Ile Glu Lys Phe Arg Ile Ala Tyr Phe Asp Lys Phe
1 5 10 15

Ser Glu Asp Ala Ile Asp Ser Phe Asn Cys Leu Ala Asn Val Ser Thr
20 25 30

Ile Ser Leu
35

<210> 37
<211> 35
<212> PRT
<213> Amino acid

<400> 37

Cys Asn Leu Thr Ile Glu Lys Phe Arg Ile Ala Tyr Phe Asp Ser Phe
1 5 10 15

Ser Lys Asp Thr Thr Asn Leu Phe Asn Gln Leu Val Asn Ile Ser Ala
20 25 30

Ile Ser Leu
35

<210> 38
<211> 34
<212> PRT
<213> Amino acid

<400> 38

Cys Lys Val Thr Ile Glu Glu Phe Arg Phe Thr Tyr Ala Asn Glu Phe
1 5 10 15

Ser Glu Asp Ile Thr Asp Phe Asp Cys Leu Ala Asn Val Ser Ala Met
20 25 30

Ser Leu

<210> 39
<211> 34
<212> PRT
<213> Amino acid

<400> 39

Cys Asn Val Ser Ile Asp Glu Phe Arg Leu Thr Tyr Ile Asn His Phe
1 5 10 15

Ser Asp Asp Ile Tyr Asn Leu Asn Cys Leu Ala Asn Ile Ser Ala Met
20 25 30

Ser Phe

<210> 40
<211> 34
<212> PRT
<213> Amino acid

<400> 40

Cys Asp Val Thr Ile Asp Glu Phe Arg Leu Thr His Thr Asn Asp Phe
1 5 10 15

Ser Asp Asp Ile Val Lys Phe His Cys Leu Ala Asn Val Ser Ala Met
20 25 30

Ser Leu